DOCUMENT RESUME

ED 100 122 EC 071 063

AUTHOR Steele, Joe M.
TITLE Me and My Environment Formative Evaluation Report 3. Design and Revision, Data Collection and Portrayal.
INSTITUTION Biological Sciences Curriculum Study, Boulder, Colo.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
REPORT NO 1520-75
PUB DATE 74
GRANT OEG-0-9-152075-3720(032)
NOTE 55p.; For additional information, see ED 087 190 and 191

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS Adolescents; Biology; *Case Studies; Curriculum Design; *Curriculum Evaluation; Ecology; *Educable Mentally Handicapped; *Environmental Influences; Exceptional Child Education; *Instructional Materials; Mentally Handicapped; Statistical Data; Student Characteristics; Teaching Methods
IDENTIFIERS Me and My Environment

ABSTRACT

The third in a series of evaluative reports on "Me and My Environment", a group-centered biological sciences program for educable mentally handicapped (EMH) adolescents, provides information about the curriculum design, the analysis and revision of curriculum materials, the gathering and processing of field test data, and a comparison of two representative seventh grade classrooms which were part of the field test. In a section on curriculum design, the rationale for selecting and organizing curriculum content is discussed; the program is traced from design to material development; and curriculum design theories and underlying assumptions are explained. Analysis and revision of curriculum materials is considered in relation to activities such as content reviews by experts and by staff members and a matrix analysis to determine emphasis on curricular components such as environmental themes and problem-solving skills. Described are problems, concerns, and costs of collecting and processing data (from field-test classes, teachers, students, and observers) to be used in evaluating curriculum effectiveness. Contrasted in a case study of the two classes are variables influencing effective use of the science curriculum such as classroom climate, student/teacher interaction, and teaching approaches that stress learning versus approaches that emphasize student behavior. Included are many photographs, charts, and graphs. (LH)

Me and My Environment

FORMATIVE EVALUATION REPORT

3

design and revision, data collection and portrayal

September 1974

Prepared by
Joe M. Steele

BIOLOGICAL SCIENCES CURRICULUM STUDY

CONTENTS

ME AND MY ENVIRONMENT

FORMATIVE EVALUATION REPORT

3

The research reported herein was performed pursuant to a grant with the Bureau of Education for the Handicapped, U. S. Office of Education, Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent the official position of the Bureau of Education for the Handicapped.

Project No. 1520-75 (Continuation)
Grant No. OEG-0-9-152075-3720(032)

Department of Health, Education, and Welfare
U. S. Office of Education
Bureau of Education for the Handicapped

the context for this report

The Curriculum

ME AND MY ENVIRONMENT is a three-year life sciences program developed specifically for 13- to 16-year-old educable mentally handicapped (EMH) children. Its development and assessment, the actual materials and their use in the classroom, the approaches to data collection, and the student outcomes all merit evaluative study. Such activities might best be viewed in the context of the three-year timeline for the development, testing, and final commercial release of the first two-year sequence (Units I-III) of the program.

So that ME AND MY ENVIRONMENT could be used in special education classes as soon as possible, the field trials were designed to overlap; two complete field tests of the materials were accomplished in three school years.

The materials for ME AND MY ENVIRONMENT include a series of teacher's manuals that suggest specific teaching strategies for each activity. Also included is a kit of all equipment, media, and supplies not usually available in special education classrooms that are required for the instructional program. There is no student text, for the program is designed around activities that are conducted by the students themselves; these activities are supported by a variety of multisensory and multimedia instructional materials, that is, materials in a number of different media designed to appeal to two or more different senses of the students. Some of these materials, in addition to science equipment in the kit, include 35 mm slides, filmstrips, cassette tapes, individual student worksheets, games, posters, study cards, wall charts, illustrated booklets, and evaluative devices. The program makes use of a 35 mm slide projector and an overhead projector. It also has tested the student's direct involvement with a Polaroid® camera and a cassette tape recorder.

The serious reader of this report will likely have reviewed, or have access to, the teacher's manuals for ME AND MY ENVIRONMENT. For that reason, information on the curriculum objectives, science content, and skill development will not be described here. (Refer to the front material in any unit of the manuals for this information.)

The development of this project and its evaluation are based upon several years of experience in developing and field-testing ME NOW, a life science curriculum for 11- to 13-year-old EMH children.[1] The ME NOW program and the first year of ME AND MY ENVIRONMENT are available from Hubbard Scientific Company, Box 105, Northbrook, Illinois 60062. Several evaluation reports on these programs have been prepared.[2]

The Evaluation

The development and evaluation of ME AND MY ENVIRONMENT began in 1971 and is expected to be completed early in 1976. (The first two-year sequence—Units I, II, and III—have been released for commercial publication.) A large number of evaluative studies have been conducted and are being utilized in the curriculum's revision. All materials except Units II and V will have been through two complete field tests and two revisions by the time of their release in their third version for commercial production.

Because the evaluation effort for this program is truly formative—an assessment designed to inform revision—the information gathered is utilized at once by the developers whenever the material that has been tested is revised. One might think, therefore, that there is little point in preparing formal reports. Yet the experience gained and some of the outcomes of the evaluation have been found to have value to other audiences. Such results are judged not only worth preserving, but also of a sufficiently general application that the timeliness of reporting becomes a concern. It seems unreasonable, therefore, to delay publication of an evaluative report until the end of the project.

---

[1] *ME NOW, LIFE SCIENCES: A SPECIAL EDUCATION PROGRAM, Biological Sciences Curriculum Study, 1972.*

[2] *James T. Robinson and Richard R. Tolman. A Formative Evaluation of ME NOW, Unit I, Digestion and Circulation. Boulder, Colorado: Biological Sciences Curriculum Study. September 1970. 97 pp. (ED 043182)*

Richard R. Tolman. A Formative Evaluation of ME NOW. Life Sciences for the Educable Mentally Handicapped, Final Report. Boulder, Colorado: Biological Sciences Curriculum Study. December 1972. 305 pp. (ED 071263)

Joe M. Steele. Arranging Field Tests: Characteristics of Sites and Students. ME AND MY ENVIRONMENT Formative Evaluation Report 1. Boulder, Colorado: Biological Sciences Curriculum Study. June 1973. 29 pp. (ED 087190)

Joe M. Steele. Assessing Student Abilities and Performance: Year 1. ME AND MY ENVIRONMENT Formative Evaluation Report 2. Boulder, Colorado: Biological Sciences Curriculum Study. December 1973. 48 pp. (ED 087191)

The volume of data being processed over the period of evaluation is considerable. Many facets of the studies provide insights that could prove of interest to an audience outside the present program developers. The studies may have significance to several specific groups, including special educators, curriculum developers, and evaluators.

In order to make the information available to such groups, and to document each major step in the overall evaluation effort, it was decided that a series of interim evaluation reports would be produced. That decision has several implications which need clarifying in order to avoid misinterpretation of the findings.

1. **The interim reports document results with experimental materials, and therefore do not present a balanced view of the program released for commercial publication.**
   For example, the reports of the measures of student performance made during the first field test do not take into account the data obtained from classroom observers, teacher reports, interviews with students, judgments of affect, and student attitudes. A report of content review and content analysis of the curriculum does not relate this information to empirical studies of effectiveness.

2. **The interim reports identify changes to be made, but do not reflect all the changes incorporated into revised materials.**
   Some aspects of the experimental program may elicit adversely critical reviews or turn out to be unworkable in the classroom. Because many readers are accustomed to seeing only *final* reports, it is possible to misunderstand the information being documented by interim evaluation studies. For example, the report of student performance relates to materials that no longer exist because of extensive revision. In no way can a judgment of the effectiveness of the final product be based on that data. Those who want summative judgments must await the final report.

3. **The interim reports do not provide a synthesis that enables one to draw a balanced judgment of the entire program.**
   It is difficult to counter the tendency to make judgments of worth on the basis of incomplete and lopsided data. A final report will synthesize and balance the totality of studies and attempt to provide an honest judgment of the total effect and value of the program. It would be premature to attempt such a synthesis in as complex a program as ME AND MY ENVIRONMENT, however, before all the data are in.

Each of the interim reports addresses concerns of interest to rather different audiences. The identity of some of these audiences is suggested by the following descriptions of the contents and purposes of the reports. The overall purposes of the series of reports are:

- –to make visible the processes, criteria, and judgments associated with the various studies undertaken.
- –to provide definitions and statements of position with respect to the new aspects of evaluation being utilized.
- –to provide documentation of the results of specific studies.
- –to provide timely reports of a number of findings related to EMH students in general, and to communicate certain aspects of the evaluation effort that may be of interest to educators in the field of special education.

A description of the completed reports concerning the formative evaluation of ME AND MY ENVIRONMENT is given below. Only a limited number of the reports were printed, but copies are on file at all special education instructional materials centers and are available from ERIC.

FORMATIVE EVALUATION REPORT 1: Arranging Field Tests: Characteristics of Sites and Students. This report serves four purposes:

1. It defines the function of evaluation in curriculum development and presents the evaluation design for the development of ME AND MY ENVIRONMENT.
2. It describes the criteria and procedures for selecting field-test participants.
3. It presents data on the actual composition and characteristics of sites and of the students who participated in the first year's field trials.
4. It reviews the placement of students in special education classes (based on the field-test sample) and draws upon the descriptions of students provided by teachers to flesh out such bare statistics as age and IQ to portray the variety of young people and instructional problems found in these classes.

FORMATIVE EVALUATION REPORT 2: Assessing Student Abilities and Performance: Year 1. The report serves five purposes:

1. It places the collection and interpretation of student data within the context of the overall evaluation design.
2. It describes procedures for the development of test items, and reports on studies of special problems.
3. It reports on the development of instruments to assess functional abilities of students, and presents results for the field-test group.
4. It reports on student performance related to instruction in the first field trials of Units I and III.
5. It presents statistical analyses of differences in performance related to a number of variables.

The Content of Report 3

The present report, **Design and Revision, Data Collection and Portrayal**, is divided into four parts:

Part I. Curriculum Design
Part II. Analysis and Revision of Materials
Part III. The Gathering and Processing of Data
Part IV. Portrayal of Reality

The articles in Part I, "Curriculum Design," do not represent an evaluation of this process. Rather, they are an effort to make the process visible, to identify the assumptions that guided the process, to describe the transitions that occurred, and to point out some weaknesses and problems that were encountered. The increasing experience of the BSCS organization in developing curricula is reflected in the progression evident in the design.

Part II, "Analysis and Revision of Materials," presents judgments of recognized authorities and staff who evaluated and critiqued the materials at various stages. The impact of these reviews on the revision of ME AND MY ENVIRONMENT, along with the field trials, was large. This section also illustrates one technique of review and as an example presents the results of a matrix analysis of Unit I. Finally, the evolution of Unit I is traced over three versions to exemplify the total impact of revision activities. Examples of prescriptions for modifying materials are presented to illustrate the kinds and degrees of change that were introduced in revision.

Part III of this interim evaluation report is titled, "The Gathering and Processing of Data." It presents the specific data collection instruments and procedures used in the testing of these curricular materials. It also describes the process by which the data generated by field trials is organized and translated into both broad guidelines for revision and specific revision suggestions.

Part IV, "Portrayal of Reality," presents a case-study view of test classrooms to examine and compare the impression suggested by various kinds of data and various sources and methods for obtaining data. ■

part 1: curriculum design

rationale for organization of curriculum content

There is no magic formula for the selection, grouping, and sequencing of content in a curriculum. The components that do become incorporated into a curriculum, however, are usually the result of an orderly approach to the problem. What follows in this and the succeeding articles in Part I is an attempt to describe in abbreviated form the considerations that led to decisions about content in ME AND MY ENVIRONMENT.

The record reflects a progression and refinement of the initial objectives and rationale for the program; it also reveals several major shifts—in the organization of content in its focus, and in the mode for delivering the curriculum to children. Throughout the development, however, certain elements have remained constant and certain procedures have operated to coordinate and synthesize the varied contributions of dozens of people. A large part of this stabilizing force is represented by key staff members whose experience in curriculum design and in the conduct of writing conferences provided control and decisive action.

At numerous points in curriculum development there are several alternatives to choose among and many decisions to make in the absence of evidence or theory. While such decisions must perforce be arbitrary, they can be informed by intuition. They can also be consistent in direction. These patterns should be apparent in the descriptions that follow.

Defining Content

The *content* of a curriculum is often misunderstood to be the choice of facts that are included in a particular subject matter area. What this report defines as content includes far more than that. Facts and concepts are one dimension, skills and processes are a second, strategies of teaching and learning are a third. A fourth dimension is the sequence of objectives and their accompanying expectations of outcome for students.

This curriculum attempts also to influence the hidden curriculum provided by the teacher and by the school—the affective climate of the classroom and the environment for learning. To whatever degree these elements are influenced by the materials, that represents a fifth content dimension. Some of these dimensions overlap. Thus a sixth dimension—the organization of learners and materials—affects the teaching strategies, the climate of the class, and the development of certain concepts studied.

As an example of how just one basic assumption can influence the content of an entire curriculum, the following consideration was introduced by the staff and advisory board early in the life of this project:

> *The major problem to be considered as the project is developed is that of educational expectancy; that is, with a range of students currently categorized as educable mentally handicapped, what educational potentials can be attained? Is training the end of any curriculum materials? Is training to develop self-respect, to be clean, to be well-groomed, and to eat certain foods to be the end of learning materials? Or can these students, in a different educational environment and with carefully designed materials, have the potential to be educated--to understand the reasons for care of the body, for example—so that they can extend their effectiveness as members of society and contribute to the quality of their lives?*

The assumption was made that these students can be and have the right to be educated, not merely trained. The educational expectancy set by the curriculum remained a problem in the field-test teachers' minds initially, as evidenced by the concerns they expressed in their feedback reports on specific activities. Some called strongly for revision of the materials to provide more training. Others questioned the educative value of certain activities. Each revision therefore did pay particular attention to this question of educational relevancy. We would be naive to expect that the materials as revised and released to the publisher completely resolve this dilemma. Thus the complex problem of creating and communicating the content continues.

THE BEGINNINGS OF CONTENT AND RATIONALE

Tracing the evolution of ME AND MY ENVIRONMENT requires a look backward at the development of ME NOW, the first of the BSCS programs for educable mentally handicapped (EMH) children, eleven to thirteen years old. When work first began on that project, meetings were held with special education teachers in four regions of the United States. An advisory board of experts in the field of special education and biology met and reviewed the needs and problems of this population. Out of the meetings came a set of basic assumptions underlying the design of the curriculum materials and a set of general objectives to guide in their development. While these have been modified slightly in form as the program evolved, the principles have served as continuing guidelines for the EMH project.

A SET OF GENERAL OBJECTIVES

1. To help the mentally handicapped child develop interests, skills, and positive attitudes through experiences with scientific—especially biological—concepts.
2. To provide the mentally handicapped child with challenging intellectual activity at a level commensurate with his ability to respond effectively.
3. To aid the child in establishing functional modes of living through heightened observation, a well-developed curiosity, an increased measure of self-confidence, and a sense of responsibility to and for his environment.
4. To contribute to the development in the child of a higher level of social maturity and emotional stability that can lead to increased vocational proficiency, realistic self-concept, creative self-expression, and more effective assimilation into the community.
5. To develop in the child a knowledge of himself in relation to his environment, along with a tendency to apply this knowledge to the tasks of everyday living.
6. To contribute to increased knowledge about the learning characteristics and limitations of the educable mentally handicapped pupil, and about effective strategies for instruction.

BASIC ASSUMPTIONS UNDERLYING THE DESIGN FOR THE CURRICULUM MATERIALS

1. Ideas must be developed with a minimum of reading on the part of the student.
2. Vocabulary, where possible, should involve *functional* rather than *technical* language, although technical names are taught when these may be useful to the student.
3. Entry points should be concerned with concrete, tangible "things," rather than with abstract, intangible ideas or concepts.
4. The classroom environment and the materials should not be cluttered with distractors; however, a variety of perceptual modes and instructional media should be used (e.g., sight, touch, smell, etc.).
5. Activities should be developed in small, discrete units that build on or reinforce a concept or skill.
6. Learning, for the EMH student, requires slower pacing, greater redundancy, and time for participation by each

student. The instructional materials should be student-doing rather than student watching.

7. An activity must involve the student in ways of applying the desired behavior; transfer cannot be assumed.
8. EMH children need, and can respond effectively to, an activity-oriented instructional approach.
9. The curriculum should be designed to provide students with an experience in science as inquiry, through the exploration of their environment.
10. Most teachers of the Educable Mentally Handicapped will need specific directions in using inquiry strategies for teaching science concepts.
11. The teachers of the Educable Mentally Handicapped, for the most part, are not science-oriented, therefore, the materials should be specific with regard to science techniques.
12. The materials and methods must permit or provide attention to individual differences and to specific learning characteristics of the population.
13. To achieve the objectives, designers of the materials should attempt to create a balance between detail and motivation; that is, the amount of minute and abstract detail that can be learned is probably a function of the interest and motivation that can be established to deal with it.

MAJOR AIMS FOR ME AND MY ENVIRONMENT

1. Development in each child of a sense of identity as a person who has some degree of control over and can act on his environment. This will lead to a degree of self-determination based on a rational coping with situations rather than on a passive compliance or an impulsive response to problems.
2. Development in each child of a success syndrome. More than anything else, each activity is intended to be a success experience for each child. It is the teacher's responsibility—almost obligation—to see that each child succeeds at a level that is challenging to his abilities and that preserves his self-respect. It is a further responsibility of the teacher to point out his achievement. As a group, the students should help each individual fit what he has done into a pattern of accomplishment.
3. Development in each child of an interest that could become a hobby or avocation over a lifetime (through an exposure to an array of experiences in science). It is hoped that many children will find some area—perhaps growing plants, caring for animals, identifying flowers, collecting things, or simply enjoying outings into the country—that they feel strongly about and can develop some competence or knowledge in. This would provide a means of self-expression and (perhaps) allow some degree of sharing or involvement with others.
4. Development in each child of a sense of relationship and empathy with other living things. It is hoped this will lead to a positive regard and caring about what affects them as individuals and as a group, because what affects them affects the community of man.
5. Development in each child of an understanding of environmental conditions that will lead to a sense of responsibility for the environment and actions that protect or improve it.

To gather information about the science concepts considered appropriate for the EMH population that the new curriculum was to serve, over fifty science syllabi in various K-12 programs of special education were reviewed by the BSCS staff and the EMH advisory board. Taking into account the interests of these children at different ages and their emerging needs for various kinds of information presented a further challenge in planning what was to become the ME NOW curriculum.

The advisory board and staff postulated that the ages eleven to thirteen represent a period in the child's physical development when he needs to understand and accept his own body. Thus, "me now" seemed a natural entry point; it fitted the egocentric perspective of the child at that age. Moreover, the sequence of biological information about the structure and function of the human body could be treated in a simplified linear manner, thus enabling objectives and goals to be established at each step in the sequence of activities.

One of the tenets of the BSCS philosophy of education is that instruction should encourage students to become directly involved in investigations, rather than to read about them or listen to the teacher tell about them. This philosophy also influenced the assumptions and strategies built into the new program.

ME NOW has proved to be a distinct departure from most special education curricula. First, there is no textbook: the student is provided with a series of activities, carefully programmed by the teacher through a manual of instructional strategies.

Second, a planned redundancy is built into the program through the use of a variety of media. Whenever possible, two or more of the student's senses are stimulated—one at a time—to reinforce the learning process. For instance, the heartbeat is *heard* through the stethoscope, the resulting pulse is *felt* in the wrist, and *seen* both in the arteries of a rabbit's ear and in a film.

Third, the principal manner of discourse in the program is through a carefully tested sequence of questions intended to minimize lecturing and to encourage student discussion and verbalization of ideas.

Fourth—and perhaps this feature of the ME NOW program is most important to an instructor with very little background in science—there is the concept of a teaching kit. In addition to specific programmed script suggestions about what to do and how to do it, the teacher is given the actual software and hardware to carry out those suggestions in the classroom.

When the ME NOW program had been successfully tested with a number of EMH classes, many of its features were carried over into the development of a new program for EMH children of junior high school age. This new curriculum came to be called ME AND MY ENVIRONMENT. Building on their experience with ME NOW, the project staff convened a conference to plan the new program in May of 1971. Participants were four special educators and one biologist who made up the advisory committee, and a writing team consisting of five special education teachers and five biology teachers. This group endorsed the focus of the program on environmental studies and suggested ways that a study of the environment could meet the needs of 13- to 16-year-old EMH students.

Following the conference, the project staff prepared an outline of the proposed content and objectives of the curriculum. In drawing up this outline, the staff made a thorough review of the physical, social, and psychological needs of adult retardates, as described in a study conducted by the Department of Special Education at Yeshiva University during the preliminary development of the Social Learning Curriculum.[1] This was a further attempt to characterize the areas of need that might be met by ME AND MY ENVIRONMENT.

As the curriculum evolved, a statement of major aims was generated. These have served as a set of ultimate goals toward which each activity should lead. ■

[1] *"Cluster Analysis of Behavioral Problems of Adult Retardates." Appendix D of A Proposal for a Research and Development Center in Curriculum for the Mentally Retarded, Herbert Goldstein, Director. New York: Curriculum Center, Yeshiva University, January 1970.*

from design to materials

THE SEQUENCE OF EVENTS

A number of approaches can be used for the production of materials: Writers can be assembled in one place or remain separated; materials can be written piecemeal over an extended time, or the writing can be concentrated within a brief period; activities can be written and pilot-tested individually, or produced in larger instructional blocks prior to tryout. The approach utilized at BSCS involves a number of writers in one location working as teams for a limited period of time to create a large amount of material. This approach has a synergistic component because of the interaction between writers and the intense, task oriented environment.

Using the BSCS model, an eight-week writing conference was held to develop the curriculum in the summer of 1971. This writing conference, composed of five pairs of biology and special education teachers plus four project staff, developed materials for more than a year of experimental use. Another eight weeks of staff review, editing, production of media, and assembly of equipment readied the curriculum for the first field trials, which were conducted from November of 1971 through June of 1972. During this period, intensive review and interpretation of processes and outcomes occurred.

A second eight-week writing conference, again utilizing ten special education and biology teachers, met during the summer of 1972. This group revised the materials tested in 1971-72 and developed a second year of materials. During the 1972-73 school year, the initial group of classes field-tested the second year of the curriculum, while a new group of classes put the revised first-year materials to the test. Final revision for commercial publication of the first-year materials was accomplished by two project staff members during the summer of 1973.

That summer, also, a seven-week writing conference composed of five pairs of biology and special education teachers and two project staff members convened to revise the materials first tested in 1972-73 and to develop a third year of new materials. First field trials of the new third-year materials were conducted from October 1973 to May 1974 in the classes that had previously tested materials for years one and two. The revised second-year materials were retested from September 1973 to June 1974 by the second field-test group. Thus, the sequence of creation, refinement and manufacture, trial, revision, trial, and final revision reoccurs for each year of instructional materials produced.

INITIATING A WRITING CONFERENCE

Once the difficult task of staffing a project, selecting and convening advisory groups, and selecting writers has been accomplished, the job of orienting personnel and initiating the actual production of materials remains. In the development of ME AND MY ENVIRONMENT, the short planning time resulted in an initial writing conference that was less structured than might be desired. Two staff members joined the project about the time the first writing conference was convened. Furthermore, sufficient lead time was not available for producing a complete plan for the curriculum ahead of time. Consequently the burden of planning and decision-making was undertaken jointly by staff and writers in the first week or so of the writing conference. Heavy reliance was placed on specialists and on studies by others in the field.

Available to the conferees were the general objectives and basic assumptions shown on page 4 of this report. All participants also received a four-page document containing some considerations related to objectives. In addition to defining three levels of performance objectives, suggesting the form for stating objectives, and outlining procedures for specifying strategies of instruction, the paper contained the following statement:

A major commitment of this project is to specify performance objectives for the EMH pupil population and to design activities through which these objectives are to be attained. The specification of objectives is influenced least by what we know about biology, and most by what we know about the EMH pupil population. We know that biology is a source of information about life and a source of understanding of the methods by which this information is obtained. But we also know that there are perceptual, emotional, and other learning difficulties as well as practical considerations for the welfare of these children in a competitive and often callous society. We seek a reasonable mix of these considerations to provide a useful, meaningful, and effective curriculum in life science for the educable mentally handicapped. We will, therefore, not pursue content to a depth greater than the ability or the need of these children to absorb it. Nor will we wish to involve these students in intellectual operations beyond their abilities to perform. These limits are still not readily identifiable and both optimism and restraint must be applied to the development of reasonable hypotheses in this regard.

The selection and statement of performance objectives for this program should represent an effort to specify (1) those content elements that represent a reasonably complete picture of environmental relationships, *(2)* cognitive and psychomotor operations the student will perform during instruction; *and, (3)* what the student will know or be able to do on his own, when an activity sequence is completed. *Individually and collectively, these three factors represent the basis upon which the effectiveness of the materials and the instruction on the learning of the child should be evaluated.*

Participants also received a twelve-page document containing a proposed outline of four units of instruction for the curriculum. For each unit, a half-page description like that shown was provided. Those areas of need in the EMH population, as identified by the Yeshiva study, were enumerated, along with possible skills, attitudes, and concepts that the curriculum could address. (See page 7.) The paper also listed cognitive behaviors to be emphasized throughout the curriculum, which included observing, describing, identifying, comparing, associating, inferring, applying, and predicting. In addition, it cited sample terminal objectives for each unit.

(Continued)

Example of writers' working outline provided at initial writing conference:

UNIT I. WHAT IS MY ENVIRONMENT?
General Instructional Objective: The student will observe, identify, measure, describe, compare, and categorize the components of his immediate environment.
Contextual Focus: Near Me
Ecological Theme: Interrelationships of environmental components.

Notes: This is planned as an introductory unit with activities of an exploratory nature to assist the student in learning to communicate about his environment, to stimulate interest in exploring the environment, and to help the student relate various environmental components to his needs, problems, and interests. Initial activities are designed to be real "grabbers."

Initial activities will likely center on the classroom and lead to observation and identification of items such as: light, temperature, air, furniture and other objects, students, teacher, sounds, smells, water, registers or radiators, dust, microscopic life, etc. The school area, path from home to school, and home can serve as the focus for additional activities designed to identify additional environment components which will be important in activities to follow: e.g., plants, animals, soil, rock, water, concrete, trash, sewage, smoke, etc.

Subsequent activities could be designed to assist the student in organizing the identified components so that he discovers some patterns, relationships, etc. These activities lay groundwork for later units, but also seem to have immediate utility in enabling the student to cope with environmental inputs.

AREAS OF NEED IN EMH POPULATION, AS IDENTIFIED BY THE YESHIVA STUDY, AND SOME CORRESPONDING SKILLS, ATTITUDES, AND CONCEPTS TO BE DEVELOPED BY ME AND MY ENVIRONMENT

PHYSICAL NEEDS

identification
sensory stimulation
body utilization
physical maintenance

Suggested Skills, Attitudes, and Concepts

The student
develops environmental awareness through sensory perception.
exhibits success in psychomotor skills.
recognizes source of dietary components.
recognizes decomposer relationships and associates these with waste disposal and body care.
recognizes sources of environmental elements to meet his needs.
chooses proper diet.

PSYCHOLOGICAL NEEDS

emotional stability
mastery
self respect
variation
self expression

Suggested Skills, Attitudes, and Concepts

The student
attends to tasks.
employs systematic problem solving techniques.
integrates basic academic learning skills.
develops feelings of confidence, competence, and self-worth.
acquires aesthetic appreciations.
senses balance, symmetry and design of environmental components.
defends position but respects opinion of others.
demonstrates modest coping and communicates about interrelationships of environmental components.
questions, investigates, and accepts his role in the environment.
persists at tasks.
initiates activities related to environment.
likes to be on work
considers alternatives.
[illegible] generally, developing
-attempts new tasks.
-weighs consequences of promiscuous sexual behavior.
Note: Many of the psychological needs should be satisfied through instructional techniques, e.g., building a success syndrome for the student.

SOCIAL NEEDS

-dependence
interaction
-social adequacy
-mobility
possessions

Suggested Skills, Attitudes, and Concepts

The student:
recognizes dependence on components of physical and biological environment.
develops receptive and expressive communication.
perceives position in space and spatial relationships.
recognizes environmental cues and utilizes these for orientation and mobility.
weighs needs for luxury items.
recognizes private ownership vs. the commons.
weighs effects of drug use on physiological function.
recognizes possible individual dependence on drugs, alcohol, and tobacco.
articulates and modifies his individual behavior with components of the environment.
modifies personal grooming and appearance habits.
values environment to extent of not vandalizing.
recognizes dependence of human society on environment.
recognizes that society has different needs and different people to fulfill those needs.
suggests, proposes, or accepts that 'no man is an island.'
investigates adult roles in society, weighs job requirements and opportunities.
weighs alternative transportation means in terms of environmental impact.
uses public transportation.
procures things that are appropriate for his environmental needs.
travels beyond former boundaries.
demonstrates appreciation and concern for environmental quality.
weighs alternative family size.
evaluates, questions, or accepts rules and customs of society as they relate to environmental issues of protection and conservation.

Within the first three days of the writing conference a rationale, objective sequence, and content outline for four units of instruction had been developed and agreed upon.

During the remainder of the writing conference, more than a full year of instructional materials were written. Included were 70 activities comprising 404 pages of printed material, plus first-draft materials related to a second year of instruction.

MODIFICATIONS IN CONTENT AND ORGANIZATION

The preceding section of Part I in this report summarized the foundations that underlie ME AND MY ENVIRONMENT and the initial decisions about its content. This rationale served as a stabilizing influence and provided a set of constants that the structure of the curriculum could be developed around. A number of changes did take place, however, and these changes had three sources. One was a series of seven critical reviews of the first year of experimental material produced. (Part II discusses these reviews in some detail.) The second source was the data from the first year's field trial of the materials. The actual data derived from field trials constitute four large volumes of material, plus two full file drawers of teacher and student responses. Examples of these data are included in Part III of this report, which explains how data for revision were organized. The third source of change was a staff review and reconceptualization of the curriculum, including a synthesis of ideas drawn from the other two sources. Portions of staff reviews that represent turning points in conceptualizing the curriculum are included in Part II.

The following paragraphs of this section describe the major changes that were introduced in ME AND MY ENVIRONMENT. The curriculum was originally conceived of as six units to be taught over a two-year period. Each unit was to be approximately twenty-five activities in length. The writers initially, however, underestimated the pace of teaching the activities and the number of activities necessary to accomplish certain objectives. To adjust the pace and retain the overall scope of the program, the activities were reorganized in the first revision to be taught in two units a year over a three-year period. When the necessary degree of expansion was provided in Unit I, however, this unit alone became a full year of activities; it was apparent that four years of materials—more than 200 activities—had been created.

A second change that was derived primarily from field-testing relates to the subject matter content itself. A curriculum dealing with environmental concerns is far more complex than one dealing with the structure and function of the human body, the subject matter of the predecessor of this curriculum, ME NOW. Because of the many interrelationships, dealing with such a complex subject matter in a simple linear manner was quite difficult. Activities and sequences of activities tended to have more than one objective, and weaving different ecological themes throughout the activities did not communicate a logical progression of ideas in sequence. To communicate the relationship and purpose of activities more clearly, they were reorganized into cores of related activities with a cluster of common objectives. This reorganization communicated to the teacher the "why intentionality" of the instructional sequence—that is, the teacher could see why particular activities had been selected, why they were sequenced in a particular way, and what the overall intent of a particular series of activities was.

Another reason for the concern about why intentionality, as well as the concern for clarifying what the curriculum is intended to do, is related to an assumption that has not been explicitly stated. The assumption is that teachers can acquire the appropriate behaviors for presenting ideas and communicating with their students through using a highly structured and detailed teacher's manual, which is the vehicle for providing teachers with an orientation toward the program. It is assumed that over time the teacher will internalize and generalize from the detailed and specific strategies presented in the materials; many of these strategies can then be carried forward to further activities and ideas that the teacher generates for the program. It is hoped as well that they will be carried forward into the teacher's classroom behavior generally throughout the rest of the school day.

While the staff believes that the foregoing assumption is warranted, it is possible that the manual does not provide a sufficient number of explicit directions for a certain proportion of the teacher population to understand and implement the curriculum as it was intended. See, for example, the case study in Part IV of this report reflecting how two of the field-test teachers used the curriculum. In any event, the assumption adds to the complexity and difficulty of developing the materials because a consistent strategy must be maintained throughout. One potential weakness in the program as a consequence of holding such an assumption may be that the teacher's manual calls little attention to teaching behaviors that could be detrimental to the teaching strategies. Thus, certain inappropriate behaviors can sometimes be maintained indefinitely by teachers attempting in good faith to use the materials as they were intended.

A third change arose because of a shift to a more explicit focus on the skills utilized in the material. The staff became aware that in the first experimental version of the material the explicit sequence of objectives related primarily to the understanding of subject-matter concepts. The writers and staff had intended, however, to place an equal amount of emphasis on the development of inquiry and problem-solving skills and applicational behaviors. The imbalance came about, it was believed, because the subject-matter content of ME AND MY ENVIRONMENT is far more complex and comprehensive than that in ME NOW. In order to assure the development of the inquiry and problem-solving skills and processes, matrices were developed to locate and plan particular skill development sequentially throughout the materials. The detailed analysis of Unit I of ME AND MY ENVIRONMENT described in Part II of this report illustrates the coordination of the many instructional components of this curriculum.

Another change that has occurred in each revision of the materials has to do with the organizing themes and sequence of activities throughout each unit. It would be quite fair to say that each writing conference should be viewed not as a mere revision of ME AND MY ENVIRONMENT, but instead as a redevelopment of the curriculum; each revision utilized a more specific and elaborate framework to extend the scope and sequence of the materials.

In addition to the modifications in the format and organization of ME AND MY ENVIRONMENT from that used in ME NOW, two changes represent the addition of components not considered in ME NOW. The first of these is the inclusion of assessment activities, called "Clues to Success," as an integral part of the materials. They provide the teacher with an immediate indication of each student's background and entering ability, and with some evidence of what the students have learned from a sequence of instruction.

The second component that has been added to the materials, "The Student Record of Progress," is a system for recording and documenting student performance across many dimensions. Built into this system is a procedure for setting different levels of expectation for different children to avoid judging all children by the same standard of success. This component is also reflected in individual activities, where a maximum and a minimum set of expected outcomes are established for each activity. ■

curriculum design considerations

THEORY AND PLANNING

To some degree curriculum development is a craft. Many curricula have been (and are) developed intuitively. Subject-matter specialists or practitioners generate the materials they feel are needed. Both Bloom[1] and Schwab[2] point out, however, that we can no longer conduct our educational activities so innocently. Robinson[3] elaborates this point: "Just as a smoker can keep smoking, but no longer be innocent of the consequences, the educated can no longer develop curricula on the sole intuitions and experiences of scholars, or teachers, or curriculum specialists, or textbook writers. Too much knowledge has accumulated in fields relevant to curriculum design to enable us to conduct our craft with innocence."

Schwab holds that five bodies of experience must be represented in the group that undertakes the task of curriculum development. These are subject-matter scholars; persons familiar with the specific target children; persons who know school, classroom, and community milieus; persons who know the particular group of teachers involved and what their skills include; and persons trained in the curriculum-making process. These need to be represented equally, according to Schwab. More emphasis on one over the others creates a distortion. To these bodies of experience should be added persons conscious of various educational ideologies and aware of their implications for the practices and goals of schooling. Finally, in Schwab's view, a period of planning is the remaining essential ingredient in the utilization of these skills in curriculum making.

(Continued)

---

[1] *Benjamin S. Bloom. 1972. Innocence in Education. School Review 80(3):1-20 (May).*

[2] *Joseph J Schwab. 1973. The Practical 3: Translation into Curriculum. School Review 81(4):501-522 (August).*

[3] *James T. Robinson. 1973. Curriculum Design. Prepared for the Annual Conference Colorado ASCD, Boulder, Colo. (November).*

curriculum design considerations *(Continued)*

How does the development of ME AND MY ENVIRONMENT fare in the light of the considerations specified by Schwab and others? Generally, staffing included persons with training and experience in the subject matter, in teaching this age group, and in working with special education teachers, parents, and children. To these were added writers who also were scholars in the subject area or teachers of the target children. The advisory committee provided some expertise in curriculum theory. The skills of the curriculum specialist and of the educational philosopher concerned with ideologies and value assumptions *per se* are rare. The project did include evaluation as a component from the beginning and in some ways, concerns about value assumptions and curriculum design were represented in the evaluation effort.

From the first, the intent of funding has been to develop a subject-matter-oriented life science curriculum for educable mentally handicapped (EMH) students. An attempt has been made to balance the training of both staff and writers in order for biologists and special educators to be equally represented. That was done as a conscious and continuing effort to respond to the needs and characteristics of the learners, and to provide concepts helpful in solving their persistent life problems. There has been an apparent tendency, in spite of this effort, for the biologists' views to carry greater weight in decision making.

Perhaps the weakest element in the project has been the lack of adequate planning time due to funding time lines. The result has been that a number of decisions were made on the basis of intuition and experience rather than on the basis of a theoretical framework. This is not a criticism of the project but of the field. The process of curriculum development in general remains vague and ambiguous. The body of theory that does exist is not adequate. Basic research has not been done.

Compounding the lack of planning time imposed by funding difficulties, staff turnover created further problems: over the course of the five-year project, thirteen persons have filled the five or six staff positions. Two persons have remained on the staff over the entire three-year period of development, however, and there was fortunately an overlapping carryover in staff from the preceding ME NOW development. Overall, a number of appropriate steps were taken by the project staff: specialists were convened to identify certain needs of the target population and areas of content that should be included in the materials to satisfy those needs. Also, a general set of guidelines was adopted. Similar guidelines had already proved effective in developing the ME NOW curriculum, which provided a successful model to follow.

ASSUMPTIONS

The design of any curriculum involves a number of assumptions, not all of them noted explicitly, and a few that go unrecognized. In this and the next section some of the assumptions of both kinds underlying ME AND MY ENVIRONMENT are identified and discussed.

First of all, it should be noted that many beliefs and commonplaces about mentally handicapped children were assumed not to be true. The project staff and the other developers of the curriculum purposely elected to explore the range and limits of ability rather than to assume it.

The effect of another assumption is that the basic design of all the material in ME AND MY ENVIRONMENT calls for activities that are built on the strategy of group instruction. This strategy has been supported by the majority of field-test teachers and has remained relatively unquestioned by the advisory board and review personnel as a classroom management technique. This does not mean, however, that children work only in classroom groups. What it means is that children do each activity at the same time in small groups. Thus, the materials are not specifically designed to enable each child to work at his own pace in moving from activity to activity, or in selecting which activities to study. Instead, the entire class is involved in one activity at a time, and the teacher must make adaptations for individual differences within the limits of that activity. It goes without saying, of course, that there may be in such activities one child who might need special help, or who might be expected to do only part of the activity, or who might be expected to perform at a lower level than other children. In such instances the curriculum is dependent for maximal effectiveness on the ability of teachers to fill in where the need for individual instruction is apparent.

Another assumption that has not been seriously questioned in the development of the program is that the curriculum should transmit a carefully selected set of information to all students. This is, of course, the assumption almost all curricula use as a starting point: that there is some set of facts, concepts, and skills that all children should acquire to some degree. The premise of behavioral objectives, criterion referencing, and competency-centered instruction is that acquiring such information and skills is the purpose of a curriculum. This curriculum effort grew initially out of a commitment to the use of behavioral objectives and a belief that certain science facts would be useful to all children. Since that time two things have occurred to modify this position somewhat.

The experience of the staff in observing field trials and in interpreting results based on the use of experimental materials suggested that, with these materials at least, learning did not always proceed in a simple linear fashion. The use of objectives and expectations of teachers and children were therefore modified.

Second, the staff began comparing the assumptions of an information-transmission model with the ideas of developmental theorists, such as Piaget and Kohlberg. Empirical evidence is accumulating that the developmental level of the child is a crucial element in decisions about content and appropriate objectives.

Clearly, the concerns expressed in this section of the report are general; they relate to all curricula. It is worthwhile to note the reach, and failing grasp, of curriculum developers. Ultimately our craft depends upon the educational theorists and philosophers. We still await an adequate theoretical and research base. Until such a time as someone can definitely say, however, "This is *the* way that children learn," a curriculum must be assessed on the basis of how effectively it can be used. The next section (along with reports of field trials) addresses this question. ■

content reviews by experts

As soon as the first experimental materials of ME AND MY ENVIRONMENT were printed in the fall of 1971, six reviewers were requested to make a content analysis of them. The materials consisted of 70 activities on 404 printed pages. The reviewers were selected to provide a variety of perspectives and were charged with performing detailed critiques based on their areas of expertise.

Edward Meyen and Warren Heiss, both authorities in the special education of retarded children and knowledgeable in the development of curricular materials, provided an analysis of the relevance of the materials to the target population of educable mentally handicapped children. Biologists Jerry Hubschman and James Jones reviewed the materials for appropriate emphasis and development of ecological themes, and for accuracy in the treatment of this subject. Biologist Edward Kormondy reviewed the materials for the adequacy and intellectual honesty of their treatment of environmental biology.

Robert Ennis, an authority in psychology and the philosophy of education, analyzed the curriculum in terms of its cognitive demands. He was asked to answer such questions as, "Can the desired inferences and conclusions be drawn from the experiences and evidence provided?" He also analyzed the development of concepts in terms of the assumptions and inferences required of the student.

The critiques these reviewers provided comprised 127 typewritten pages. Their detailed suggestions for specific changes in many of the activities proved invaluable for revision. Excerpts from some of their general and summary statements are presented here to indicate the flavor of their reactions. These criticisms reflect the kinds of problems that have since been addressed in two revisions.

The special educators who reviewed the materials had been acquainted with the previously developed ME NOW curriculum. Both felt that ME AND MY ENVIRONMENT represented an improvement over the earlier program. Warren Heiss made the following comparison of the two curricula:

> *There is a major shift in content from the ME NOW materials to the ME AND MY ENVIRONMENT units. This shift has caused a change in the concept and activity flow. These new materials are more like a loose collection of activities from an integrated unit. This is a relative statement. It does not imply that there is no integration, but in contrast to the ME NOW materials there is a looseness.*

> *One reason for this looseness is related to the notion that the point of reference for study, the environment, is less easy to define than the content of ME NOW as a referent. The resolution of problems presented in the ME AND MY ENVIRONMENT materials are less finite and require more discussion. One way to combat this problem is to develop modular activities for teachers to follow with sample options for the students to pursue.*

The "looseness" was also noted by other reviewers in terms of the sequence and flow of activities, the problem of transition or of integrating some activities with others, and the conceptual shifts that occurred from time to time. Teachers, too, indicated some difficulty at times in understanding the relationship and progression of activities.

The label "why intentionality" came to be applied to the problem of clearly communicating a progression of ideas and a purposeful flow of activities. These concerns led to the reorganization of materials into Clusters of Related Experiences (CORES) prefaced by overviews for the teachers.

Other concerns raised especially by the special educators were questions of the relevance and application of some of the activities. Edward Meyen elaborated on this point:

> *I do feel that the activities are highly impersonal and they do not provide any social frame of reference for the child. In stating this, I am not implying we cannot teach these particular lessons to children. I am sure in most cases you will be successful. I am just making the point that you need to capitalize on this by also at the same time allowing the child to relate these concepts to his social frame of reference, which is life in general. I have the feeling as i review the activities that you are trying to maintain a great deal of objectivity in that you seldom require the person to refer to himself and his interaction with the elements of his environment. The emphasis is generally on "what's about me" rather than a personal perspective.*

Comments of the biologists reviewing the materials were primarily directed at the subject matter. All felt, as Jerry Hubschman stated, "that the terminal objectives outlined are quite realistic. In most cases the subordinate objectives represent a pretty good fit to the higher categories."

Several noted inconsistencies in the level of sophistication expected from students. An example of this was use of the word "calorimeter" but substitution of "germ jello" for the term "agar."

(Continued)

James Jones expressed this concern:

Many of the activities are planned as though science were the only subject being taught. Not only does the time involved in some activities seem extensive, but each activity is played for its environment content. I realize that many teachers will bridge the gap between disciplines, but for others it may be necessary to integrate some of the major areas - at least in a general way.

Edward Kormondy summed up the treatment of content by saying:

Sufficient information is provided to convey and/or develop basic ecological understanding of the operation of key processes such as the flow of energy and the movement of nutrients and of the interacting role of organisms and environment in effecting and affecting those processes. That which is developed is sound and fundamental to appreciating the overall objectives of the program, and particularly to enable the youngster to come to some degree of understanding of "him" with respect to "his environment." ...Considering the overall objectives and the nature of the clientele, I believe the "coverage" of environmental elements and problems is amazingly complete...Some of the objectives and concepts (save one on personal hygiene) are trivial in the ecological framework. Some are less significant to the "big" picture, but in terms of personal relevance for these youngsters and most others at that age, they are important and nontrivial. The activities are "hands-on" and this is good and hence not trivial; further, and more importantly, they lead somewhere.

The educational psychologist, Robert Ennis, summarized his remarks on the logical demands of the curriculum in this way:

In general I find [the activities] to be an exciting package, and in retrospect think that in my comments I have emphasized six ideas:

1. *That there be greater emphasis on the need for controls in experiments.*
2. *That students be more involved in planning the experiments.*
3. *That they be asked the question "Why?" more frequently.*
4. *That in planning experiments the strategy of imagining possible results and deciding what one could conclude from those results be used frequently.*
5. *That the introduction of the term "environment" and the technical term "work" be reexamined. I am not urging that these be eliminated. I don't know enough about the abilities of the students under consideration. But the introduction of these terms, especially the technical term "work," will cause some problems, and even if successful, will not result in much significant learning.*
6. *There were a number of times when inferences were drawn without an adequate basis.*

All of the specific criticisms and suggestions contained in the six reviews were sorted by activity or by related clusters of activities for use by the revision team during the summer writing conference. Complete copies of the six reviews were sent to all advisory committee members, reviewers, and project staff prior to a meeting of those persons in January of 1972. In preparation for this conference, project staff conducted a thorough analysis of the reviews and summarized the major ideas, as shown in the inset.

In addition to considering this summary, conferees discussed the overall organization of the curriculum, as well as a number of problems concerning the developmental time line; they reviewed the formative evaluation design, examined the teacher education aspects of the curriculum implementation, and discussed the development of guidelines for the revision writing conference. Major outcomes of the conference included:

1. The suggestion that the project be extended to develop a three-year curriculum in order to encompass more adequately the environmental concepts and concerns that are particularly germane to such a program.

2. Suggestions for reorganization of the curriculum that would overcome many of the problems reflected in Part I of the summary (see inset).

3. Resolution of the questions raised in Part II of the summary (see inset).

4. Suggestion that the project staff undertake a series of matrix analyses of Units I and II of the materials to provide a further basis for developing guidelines for revision.

SUMMARY OF MAJOR IDEAS USED IN THE REVISION OF THE EXPERIMENTAL EDITION OF ME AND MY ENVIRONMENT

Part 1. Ideas, suggestions, and criticisms that represent consensus opinions because they appear in several reviews.

a. The curriculum is successful in creating an effective mode of presenting a lesson to the teacher The format, instructions to the teacher, guidelines for pupil-teacher dialog, and the specification of resources are usabie.

b. The terminal objectives are realistic and, in most cases, the subordinate objectives and activities should enable the former to be achieved.

c. None of the objectives, with the possible exception of the emphasis on personal hygiene, are trivial in an ecological framework.

d. The activities are "hands-on," usually relevant, and usually lead somewhere.

e. The overall conception is solid and the implementation is consonant.

f. There are large inconsistencies in the level of sophistication required of the student, both in terms of vocabulary and of anticipated cognitive behavior.

g. In many cases there is a lack of continuity between activities.

h. More effort is needed to provide the teacher with means and entry points for integration into other curriculum areas.

i. More effort is needed to help the student relate the activities to himself and his interaction with the environment, and to provide application of the knowledge and skills to environmental questions and problems that the student faces.

j. Greater emphasis should be placed on the role of microorganisms in the cycling of materials.

k. Implications for teacher education created by the inquiry-oriented approach need further consideration.

l. An ecological theme, "Ecological Trade-Offs," should be added to the curriculum development model. Activities that deal with both ecological concepts and environmental problems should place more emphasis on consequence.

m. Too much emphasis tends to be placed on the use of indicator solutions and chemical tests when the same concept could be developed by using living organisms or natural systems.

n. Students are asked to overgeneralize and to draw inferences without an adequate data base.

o. The students are asked to place blind faith in instrumentation. Activity sequences should be developed to test and verify the accuracy of instrumentation.

p. The methodology of science needs greater emphasis; an understanding of the methodology of science should be more logically developed to include the need for controls in experimentation, the interpretation of data, and the involvement of students in experimental design.

q. The question "Why?" should be posed to students more frequently.

Part 2. Ideas, suggestions, and criticisms that require further discussion and resolution.

a. Several reviewers expressed concern about the choice and use of several vocabulary words and related concepts. These include:
   1. Definition of the term "environment"
   2. Incorporation of the technical notion of "work"
   3. Use of the term "germ" in the context of both beneficial and harmful microbes because of the negative connotation usually ascribed to that term
   4. Use of the technical definitions of "food" and "nutrient"

b. Consideration should be given to adoption of metric units of measurement.

c. Several reviewers expressed concern regarding the possible collection, culture, and spread of pathogens in some of the Unit II activities.

d. Less emphasis should be placed on personal hygiene, and those activities concerned with it might be made more sophisticated.

e. Greater emphasis might be placed on the role of the oceans in food supply, oxygen supply, water and nutrient cycling, and as sewage basins.

f. Objectives should not be couched in strict behavioral terms, which in many cases require only memorization or recall of appropriate responses, but should be stated more generally as "know," "realize," or "understand" kinds of objectives.

g. Are various types of scientific apparatus included because of their novelty, and would the activities be more relevant if some of this apparatus were replaced with more common household instruments?

goal free evaluation

As part of his work on current problems in evaluation for the National Institute of Education Evaluation Planning Team, Michael Scriven suggested that a useful supplement to formative evaluation would be a "goal free evaluation" conducted by an independent evaluator. Such an evaluator would have the objectivity to note unintended side effects as well as unrecognized positive effects of the program. In an additional attempt to ascertain if the curriculum goals were being met, the project staff decided on such an approach as a valuable complement to the evaluation of ME AND MY ENVIRONMENT. Arrangements were made with Ernest R. House of the Center for Instructional Research in Curriculum Evaluation (CIRCE) at the University of Illinois to be the independent evaluator. He conferred with Michael Scriven and with Robert Stake and Thomas Hastings of CIRCE to refine and adapt the technique to this project. Donald Hogben, a visiting scholar at CIRCE, worked with him in reviewing the written materials. The evaluation was accomplished in February, 1972. The project staff found his efforts insightful, containing many suggestions that have been utilized in the revision and further evaluation of this curriculum. The remainder of this article represents his written report of the study.

Readers will, of course, recognize that the following represents the opinions and reactions of one man. His qualifications are these: broad experience in the evaluation of educational programs in all subject areas at every level from preschool through high school; experience in inquiry teaching and evaluating this instructional approach; experience in curriculum development; and a working knowledge of many of the newer techniques in evaluation.

House does not, however, have a background in biology. He chose to assist him a person trained in biology and acquainted with BSCS materials. Neither, however, had experience with EMH children. Perhaps some insights were lost because of this. However, the payoff of using a goal free approach was high, as reflected in the article which follows.

A GOAL FREE EVALUATION FOR THE FIRST EXPERIMENTAL EDITION OF ME AND MY ENVIRONMENT

Scriven has recently suggested that the evaluator should deliberately ignore the goals of a project during evaluation. This would prevent the "intentional fallacy," where the evaluator attends to the intentions of the program designers and ignores important effects the program might be having regardless of the expressed goals. It would allow the evaluator to concentrate on the effects of the program (which might be either better or worse than the designer intended) and prevent premature closure. The initial phase of such a design was carried out with the encouragement of the Biological Sciences Curriculum Study (BSCS) during the first field trials of ME AND MY ENVIRONMENT, a biology curriculum for 13- to 15-year-old educable mentally handicapped (EMH) students. Such an exercise was intended to give the BSCS developers of the program a new perspective on their curriculum and suggest new emphases for their formative evaluation.

PART I. THE PREVISIT HYPOTHESES

Prior to the visit the evaluator (Ernest R. House) spent about four days reviewing the previsit materials, which included:

A. Two units of ME AND MY ENVIRONMENT (not reading the goals)

B. Six reviews of Unit I by various experts

C. A three-day sequence of observation reports from one classroom and the Pretest for Unit I
Information on the Pretest (distribution of scores and item correlations)
A copy of the teacher feedback form
The evaluation design, time schedule, and evaluation questions

Reaction to Unit I Experimental Materials (House)

1. Materials may be viewed as a positive good since they fill a void where nothing previously existed.

2. Do mentally retarded youngsters need to do experiments and the kind of problem-solving suggested by the materials? Are units of measurement really necessary? The materials are heavy with basic concepts of biology. No doubt they are favorably viewed by biologists. There seems to be less concern for what these boys and girls might need to know. While the "content validity" of the materials would seem to be sound, the "utility validity" is open to question. (This might suggest a goal priority study of the materials at some future date.)

3. The materials read very well and would be interesting to the teacher, I suspect. This would also enhance their

marketability. The materials also maintain the central role of the teacher and teacher "performance" as a central concept. (This could be tested by items from the teacher response form.)

4. The materials are likely to dramatize the lessons, e.g., plants need air to live. However, they are highly unlikely to develop "problem-solving" skills, "scientific method," etc. By mechanically following steps in experimental procedures, students are unlikely to generalize to other situations. (See Wittrock's study of inductive teaching.) The sequence of materials, i.e., building on them in later lessons, mitigates against the drawing of conclusions because the time spans are so great. (This hypothesis is feasible by testing but not with the items on the pretest.)

5. The inquiry teaching approach has been adopted throughout the materials, which means that economy of presentation has been forfeited. It would seem that giving up efficiency must be balanced by gains in either (1) fun in learning, (2) greater understanding, (3) greater transfer, (4) a more active rather than passive role for the students, or (5) some combination of these. Both (1) and (4) seem to be rather important for these students. It would seem that giving up such economy should result in tangible gains. One must also be sure every principle is extremely important.

6. Even if gains in items (3) and (4) are found, the materials might be helped by providing a few hints at the beginning of the activity that would be extremely helpful to the students. Even Bruner has said that acquiring a skill requires that the goal must be plain. One must have a sense of direction. Since the exercise of any skill is governed by intention and feedback, one must make the purpose clear. If one does not know the goals of these lessons, their purpose is often unclear even to a highly sophisticated audience. The mentally retarded youngsters must really be lost much of the time. Too much of this produces hopelessness rather than inquiry. I believe the authors of the materials have not penetrated the phenomenological world of the student. It is conceivable that these materials could make these boys and girls more dependent.

7. Teaching inductively is often dependent on the verbal quickness of the students. Otherwise, asking questions is like pulling teeth. Again, classroom observation and teacher feedback might answer this question. Also, variation in presentation modes is worth thinking about.

Summary: Considering just the materials themselves, the program seems very attractive. This is especially true considering that no competition exists. The materials would seem to be appealing to teachers and quite strong in biological content. The possible weakness lies in the exclusive reliance on the "inquiry-inductive" mode of teaching. Some of these pedagogical assumptions have been outlined. Some of those are subject to empirical (but not easy) verification. The longer-range questions of how much good this material will do the student in later life was not discussed.

Review of Six Critiques (House)

The critiques by six experts tend to confirm my own thoughts about the materials (since there are numerous suggestions on individual activities, I will confine my comments to overall trends I see):

1. The content is very well covered and appropriate, suggestions for improvement of individual lessons notwithstanding. The reviewers also concur in thinking the materials very good.

2. The sequencing of activities is suspect. Frequently the reviewers mention "a common thread needed," the lack of integration of the activities, and the difficulty of shifting to new materials. Unevenness in presentation and sophistication of the materials is mentioned.

3. Perhaps associated with the integration of activities is questioning the "relevancy" of activities. One reviewer in particular talks about the "impersonal" nature of the materials, which I also felt. There is a heavy reliance on "science" as such and lab techniques, with little immediate concern for application. Using experiences and personal things from the students' lives would enhance applying those concepts as well as their relevance. Surely with these youngsters, applications should be as important as understanding.

4. Do the students get the point of some of these activities and do these exercises faithfully reflect the underlying concept? Numerous cases of specious reasoning also call into question the suitability of various exercises for enhancing reasoning abilities.

5. Teacher utility of the materials is extremely high. They are quite attractive materials. The tendency of teachers to turn the inquiry teaching into a guessing game is noted.

Evaluation Materials

It is clear from the Observer's Notes,* though no great surprise, that the teachers do not follow the materials precisely. It is doubtful that any amount of training could eliminate such deviance, and the materials will have to operate under conditions of low fidelity. It would be worthwhile to examine a longer sequence of observer notes at the BSCS office to look for systematic digressions.

The pretest items and subscores are not highly correlated. Is the test measuring what it proposes? The best predictor for total score is the Piaget conservation of liquid item. Are some students operating at one logical level and some at another?

The evaluation design is more like a quality control rather than a testing of the basic assumptions, which is perhaps appropriate for a formative design. The evaluation questions and priorities are good. The classroom observation is a particularly strong feature of the evaluation. Overall, the evaluation is very well thought out, thorough, and probably overly ambitious.

PART II. ON-SITE VISIT

The two-day on-site visit included:

A. Observing two classes, talking to several students and both teachers, and talking to one classroom observer.

B. Going through observer notes, files, and teacher reaction forms in BSCS, and reading student interviews from other sites.

C. Interviewing major staff members and the BSCS associate director about "success" of the project and eliciting criteria of success from each.

D. Giving staff a preview of conclusions.

Classroom Visits

1. First class—a middle-class suburb

Students are enthusiastic and having fun, but cannot follow logic of lesson. The students have fun manipulating materials, moving around, etc. Teacher has usual problems in question session since kids are not following. The teacher cues on correct hypothesis, ignores false.

On their own, students really don't know how to design experiment, but they have fun. Some believe that plants can

(Continued)

**In field tests, four classes were observed by the BSCS project staff during each science period.*

live with air or sunshine. Others don't have any idea why experiments work or don't work. Many like the class, however, because they get to "fool around."

2. Second class—inner city, almost all black

Totally different--total apathy. Students are drawing maps but no one knows how to do it, nor does anyone care. They say they don't like class. Students in the other class were motivated; these are not. Different teacher, activity, and students. The observer thinks the main difference is in teacher style. The teacher and the school are afraid the students will get out of hand. She resorts to making students copy materials when they get too excited. Whatever the cause, the materials are clearly not working here. Talk with teacher confirms school discipline is the problem. Nonetheless, she says she really likes the materials.

BSCS Files

By the next morning, I am already set in ideas. Reading through the observer notes, teacher reaction forms, and even student interviews only confirms the hypotheses I have. By this time I am not open to new ideas. [Note: The evaluator is suggesting at this point that he has acquired a framework or "set of prejudices" about the project such that any further data reviewed tend to be interpreted from that framework. Hence he is no longer "goal free." As the intent of this procedure was to arrive at a framework independent of the ideology of the project, the evaluator is indicating that such an outsider's view was fully formulated and tended to become a closed view by this point. Thus, further data, rather than contributing to the formation of a point of view, is merely interpreted according to a point of view. Further study thus becomes self-defeating, as the major insights of a new perspective have now reached a point of closure similar to those of the project staff.]

Staff Interviews

Finally, the major staff members, including a BSCS administrator, were interviewed to ascertain their global impressions of the success of the project, and especially the criteria for making the judgment. All thought the project would be successful because it was filling a vacuum—no other science materials for EMH students exist. Below are the criteria for "success" (in order of importance) as seen by the project staff:

(Associate Director)	(Project Director)
(X) 1. Influences other projects	(X) 1. Students treated well, "normal"
2. Produces materials	2. Students learn new material
3. Accepted by teachers	
4. Involves the students	
(Science Consultant)	**(Science Consultant)**
(X) 1. Accepted by teachers	1. Accepted by teachers
2. Students like material	2. Students like material
3. Students do better on tests and teacher ratings	(X) 3. Students understand science concepts

The goal priorities for the project are rather similar. Producing materials and getting them accepted by teachers is the highest priority of all. This is a reflection of the strong "production" orientation of the organization, often expressed in the logic, "Unless teachers get the materials and use them, there is no chance of students improving." There is no confusion about the overall mission of the organization. However, even though there is considerable similarity in overall goal rankings, there are significant differences in the emphasis each individual gives certain goals. Strong emphasis on a particular criteria relative to the others is indicated by an X. These differences are manifested in various ways in the project.

PART III. FINAL HYPOTHESES

1. The materials seemingly replace not much of anything, and hence must be considered a positive good.
2. Teacher satisfaction with the materials is quite high, even where they seem not to work too well. Giving teachers sets of activities and materials they can use strikes a responsive chord.
3. The scientific validity of the materials is high.
4. The inquiry approach:
   a. results in fun and interest among the students. Combined with the "hands-on materials" and lab activities, student involvement seems to be the main outcome of the curriculum,
   b. results in a dramatization of facts, e.g., "plants need light to grow," likely to result in long retention,
   c. does not result in higher cognitive discoveries such as the "scientific method" or the inductive mode of thinking. Outlining the steps in a scientific reasoning process does not make one think "scientifically" any more than working with syllogisms makes one logical. The method by which scientists learn to think in their research paradigm is quite different, modeled by the graduate school. In fact, if teaching "scientific method" was the goal, the goal itself is of dubious value.
5. The relevance and sequencing of many activities is questionable. Many activities make sense only if they are related to others across long sequences of time. It is doubtful if students can make these connections.
6. The purpose of many activities is unclear from the student's viewpoint. He often does not know why he is doing something. It would be helpful to introduce some "intentionality" into the activities by briefly providing some sense of direction at the beginning. Again, even Bruner has said that acquiring a skill requires a sense of direction. The purpose must be clear. A few of the activities might be rewritten along this line and tried out, still maintaining the basic inquiry approach.
7. The black students do not seem to be of the same type as the white EMH. They lack interest. It might be desirable to create special materials for those students. The observers, particularly Ray Fisher, might be used to help revise the materials and make them more relevant to the inner city. Since any "future" orientation is lacking, it is all the more important that the activities be immediately relevant.
8. There is no question that the materials will not be taught as the designers intended. One condition of use is that they will be misused to varying degrees.
9. It might be useful for the designers to teach the materials themselves. Certainly some insights will be generated that are not available through documentation. Generally the evaluation has worked at a fairly literal level. That has been successful at changing specific points in the material. Another possibility would be for the evaluation to inform insight a little more and reduce the paper burden. ■

staff reviews

"EXCERPTS FROM JUNE 1972 REVIEW"

A number of staff reviews occur as each component of the curriculum is developed, tested, or revised. Recurring studies include the review of feedback from field trials, summary and discussion of external reviews, and matrix analysis of curriculum structure. The latter is described in the following article.

Prior to each intensive revision effort, a comprehensive staff review is completed. The first such statement was produced prior to convening the second summer writing conference for the project in June of 1972. This report included a rationale for revising each of the tested units and suggestions for overall revision of ME AND MY ENVIRONMENT. The report also included a review and reorganization of the program objectives, a paragraph summary of the analysis of each activity, and specific suggestions for the way it should be revised.

The following excerpts from this report represent the views of the project staff in June 1972. This synthesis of reviews, feedback, and observations was based on the first version of the materials. Subsequent revisions have corrected the defects described. The staff perspective is correspondingly positive. Most of the recommendations in this review were taken into account during the revision of the experimental materials, greatly enhancing their effectiveness. Because the statement does reflect a major developmental step in the production of the curriculum it is included here.

In the experimental materials the nature and format of the objectives appear to be restrictive, tend to stifle the creativity of both writers and teachers, and in many cases are either vague or unrealistic. Meyen, in his review, questioned the supposition that a hierarchy of objectives exists; he suggested that there are many pathways to achieve an objective. Ennis criticized the language used in stating objectives, pointing out that some are extremely vague, providing little guidance to the teacher, while others are so specific as to invite meaningless learning and evaluation.

To overcome these problems partially, we might follow some of the ideas suggested by the Science 5/13 project, as outlined in With Objectives in Mind.[1] *First, if we are going to use objectives at all, we need to develop a way to embed them firmly in the teacher's mind. The present format does not seem to accomplish this. Perhaps that could be corrected in part by including a detailed discussion of the use of objectives in the introductory material. Secondly, we should eliminate the terminal objectives, and, in their place, list the broad general aims of the unit at the beginning of the unit. Then, we should break the unit into activity clusters. For each cluster, we should list the broad aims which that cluster might lead toward. In addition, we should list the more specific objectives of that cluster, followed by a discussion of how the activities might enable students to reach the objectives. For each activity, we should state a purpose that would include how the activity should get at various objectives, and list the operational ("student should") objectives.*

The goal-free evaluator may well be right that we have carried the idea of inquiry to the extreme and that children are kept in the dark most of the time. They simply do not understand the purpose for doing many of the things that we ask them to do and, therefore, may not be very interested in doing those things. House said, "The mentally retarded youngsters must really be lost much of the time. Too much of this produces hopelessness rather than inquiry. I believe the authors of the materials have not penetrated the phenomenological world of the student. It is conceivable that these materials could make these boys and girls more dependent."

Three major thrusts need to be made in the 1972 revision. The first is to give the student direction and a sense of purpose. The second is to provide continuity between activities and clusters of activities. And, third, we must get to a consideration of consequences and application.

Because we have found performing a matrix analysis for Unit I to be a revealing (albeit laborious) task, we should place a priority on developing a similar checklist for the writers to follow and complete as the activities are revised or developed. Since it is improbable that a writer can keep the overall curriculum model, the feedback and other data, the reviews, etc., in his head, a staff member should be charged with the responsibility of reviewing activities with as many of these things in mind as possible. (This would require continued review of ALL the data.) He could constantly remind writers of things they are overlooking in an attempt to insure that they pay attention to the guidelines, suggestions, feedback, etc.

(Continued)

---

[1] *Len Ennever and Wynne Harlen. September, 1971. With Objectives in Mind. University of Bristol School of Education, Schools Council Publications. London: Macdonald & Co.*

Unit I is still seen as primarily an introductory package which helps the student develop a concept of environment, gets him interested in looking at the environment, gives him purpose for studying the environment, helps him identify components and deal with environmental inputs through categorization, and makes him realize that he and all other life are dependent upon things that the environment provides. Attempting to reorganize the unit has "made one thing perfectly clear": we can't do everything in the first unit! What we have to do is build the background necessary to do it all in the units that follow. Therefore, as House pointed out, we have to weigh each activity very carefully to insure that it is important because we can't waste time on things that are not.

Contrary to earlier opinions, we should probably not put a lot of emphasis on experiments, experimental design, and the like. We know that students are intrigued by the equipment, and they should have the opportunity to mess around with this "science stuff" early in the unit. We should be careful, however, not to jump into the deep water. As pointed out in With Objectives in Mind, *a child who has missed early experiences of a particular kind needed to develop the concepts required for making sense of later experiences in science is not likely to be in a position to benefit from these later experiences. "We do not want children to 'mark time' on activities which are insufficiently challenging for them; but the opposite mistake is probably more common and damaging, when children are introduced to ideas or expected to deal with problems for which they are not ready,... [this] can lead to children losing interest or uncomprehendingly following instructions." (p. 17.)*

House said, "[that he sees] the exercises as being a dramatization of the facts and I really can't see them leading to scientific thinking." Activities that emphasize logical processes, which we usually refer to as "scientific method," need to be simple, clear-cut, and fail-safe, especially in Unit I. The present "plant need" experiments are not this, for example, and reading the feedback or interviews will quickly support the idea that most of these students are "uncomprehendingly following instructions." We can probably assume that: these students will find difficulty in separating the effects of two or more variables; they will not be systematic in combining variables except in the simplest situations; they will be content with describing rather than explaining results; and they are unlikely to abstract generalizations from the results of the demonstration kinds of experiments.

It is suggested that Unit I should emphasize observation, description, identification, speculation and guessing, comparing, some associating, and a lot of value judging. Asking questions and experimenting (in the sense of "messing around" with things) should be emphasized. Opportunity exists for introduction of the idea of variables and controls, and these might be introduced in very simple situations and then reinforced every time the opportunity to experiment comes up. We should identify what we see as a logical sequence of introducing ideas and concepts of "science method," build a check sheet and score activities to keep track of what is introduced when, and then (for revision of Units I and II) specify for the writers *where a certain idea is introduced and where it should be reinforced.*

Another concern which the staff should thoroughly consider is that many of these teachers, in general, do not understand enough about science to teach "scientific method." In fact, the present curriculum may do more harm in creating wrong impressions about science than it does good—and the teacher variable is significant here. We can't assume, as writers (and staff) have in the past, that the teachers will understand experimental design, variability of results, the meaning of inference, extrapolation from data, generalization (and overgeneralization), etc. Thus, it is important to build into the strategies, bit by bit, these sorts of things for the teachers to learn right along with the students! WE NEED TO DEVELOP A PLAN FOR THE ENTIRE CURRICULUM WHICH WILL ACCOMPLISH THIS—WHICH CAN BE SUPERIMPOSED ON THE ACTIVITY SEQUENCE.

Finally, there are some areas that conferees at the January meeting suggested be included in Unit I, but which we have not been able to fit in for one reason or another. First, it was suggested that activities which include the child as part of a social environment (part of a population) be developed. The VD activities get at this to a small extent, and in a somewhat dubious way. This seems questionable as a realistic goal for Unit I. Second, it was suggested that something be included that gets at the idea that man's needs lead to exploitation of resources and interferes with natural processes of ecosystems. This also seems unreal for a Unit I goal. Similarly, the idea that resources are finite and life depends upon recycling of resources seems to be an unrealistic goal to develop in the first unit. We can lay some groundwork (and should), but this is not an objective of the first unit.

CONTINUED REVIEW OF MATERIALS

In the interval since the reviews and judgments above were made, all of the units of ME AND MY ENVIRONMENT have been revised and further field testing of the revised versions has taken place. Several reviewers were asked to critique the revised materials, and their judgments are being incorporated in yet another revision that also draws on field-test results. The following statement reflects the view of a special educator, Edward Meyen, about the revised materials:

"I think the activities in terms of explicitness, relationship of activities to objectives, and the teacher information are much improved. My major concern is that each core and in many cases individual activities are presently in reference to rather isolated concepts. Your concern appears to be primarily for teaching the content (i.e., science and environmental influences) without due consideration being given to the learner. I would still like to see more use made of either a theme or some frame of reference which is extremely meaningful to the student and from which you can generate a frame of reference for teaching about the environment. To teach specifically about temperature, habitats, and magnets, etc., means that the student must work backward to full understanding of his relationship to his environment. I question that this is necessary and sincerely feel you could approach this in such a way that the need for information on habitat, temperature, etc., could evolve from a broader orientation that you could build on the part of the student."

The first two years of ME AND MY ENVIRONMENT has now been released for commercial distribution. A review of this final product should be commissioned to provide a judgment of the materials from a somewhat different perspective. The purpose of the review would not be primarily to aid revision (although it will have implications for units still under development). Instead the curriculum would be judged for its relevance and utility as a component of special educational programs. The problems and promise of introducing the program into schools would be studied. Concerns for program dissemination and implementation would be raised. ■

a matrix analysis of the curriculum

Part I of this report alludes to a number of considerations that directly affected the content of ME AND MY ENVIRONMENT. Many of these components were formally defined during a matrix analysis of the materials that was first undertaken by the staff in the spring of 1972.

A matrix analysis is a process of determining the frequency and degree of emphasis on specific curricular components that occur and reoccur throughout the program. A display of this information similar to a graph allows one to inspect the sustained development of each component across activities. It also allows one to study the number and variety of components emphasized within each activity. In this way, the complexity and scope of the program can be made more visible. In this article five of these displays, or matrices, appear as tables. Each matrix is related to the broad general headings of Environmental Themes, Inquiry Skills, Other Developmental Abilities, Problem-Solving Skills, and Variety of Student Involvement. Because ME AND MY ENVIRONMENT is an activity-oriented curriculum, a number of categories relate to opportunities for participatory experiences in groups or as an individual.

At the time of the first matrix analysis, a separate study was made to determine the density, repetition, and development of terms, ideas, and concepts in the materials. The results of this study were not formalized in a report but were incorporated in the plans for revision of materials.

The first matrix analysis was carried out activity by activity. Each category was rated in four degrees of emphasis: a) whether the idea or concept was a central feature stressed in the activity; b) whether it was present, but unstressed on the assumption that students could assimilate it without elaboration; c) whether it was suggested or implied, but undeveloped in the activity; or d) whether it was not present, but needed for the activity.

As the matrices of categories were applied to the experimental materials, the potential utility of this kind of analysis as a tool for development and revision became apparent. In one sense, such matrices constitute a curriculum model that graphically displays a complex scheme for program design. They could be used not only to identify, but also to specify the emphasis for each activity.

As a result of the analysis, a charge was given to the 1972 summer writers to use the matrices of skills, themes, and concepts in their revision or creation of materials. This experiment in the use of the matrices, however, proved to be far too time-consuming and difficult a task for the writers to accomplish during a summer writing conference. But the matrices did serve to orient the writers to the multiple purposes of the curriculum.

The extensive revision of the materials during the summer of 1972 necessitated a completely new matrix analysis, and this was later accomplished by the staff. Again, the analysis provided some of the direction for the next revision. As the variety of data obtained on each activity was summarized, these matrix categories entered into the prescriptions for needed revisions that were developed to guide writers in the 1973 summer writing conference.

The most valuable outcome of the rather laborious task of analysis was the clarification it brought to the inclusion and development in the materials of inquiry and problem-solving skills. In the initial version of ME AND MY ENVIRONMENT, the inclusion and placement of these skills was intuitive. The matrix analysis resulted in placing far more systematic emphasis on creating materials that further the development of these skills. The intentions of the project staff came to be these: (1) emphasis in the materials should be divided equally between content and the two groups of skills; (2) within each year of materials, there should be a sustained emphasis on the development of each skill, and the more difficult or complex skills should be introduced one by one, after the simpler skills have been developed; (3) the materials should not attempt to develop these skills sequentially over the three-year period, but should redevelop them each year; (4) the simple skills should be extensively developed in the first portion of the program, while the more complex skills should be developed as extensively as possible in succeeding years.

How fully these intentions have been realized in the final product cannot be determined until all the units have been revised and released for commercial publication. An *indication* of the pattern of emphasis can now be obtained, however, inasmuch as the first two years of materials has been released for commercial distribution. The matrix analysis of Unit I (commercial edition) of ME AND MY ENVIRONMENT follows as an example of what was done. The pattern of emphasis will be discussed in terms of five broad dimensions of skills or conditions. Refer to the front matter of any unit of ME AND MY ENVIRONMENT for the definitions of categories used in the analysis. The categories of emphasis under each of the main headings were coded as either central, present, implied, or absent in each one of the activities.

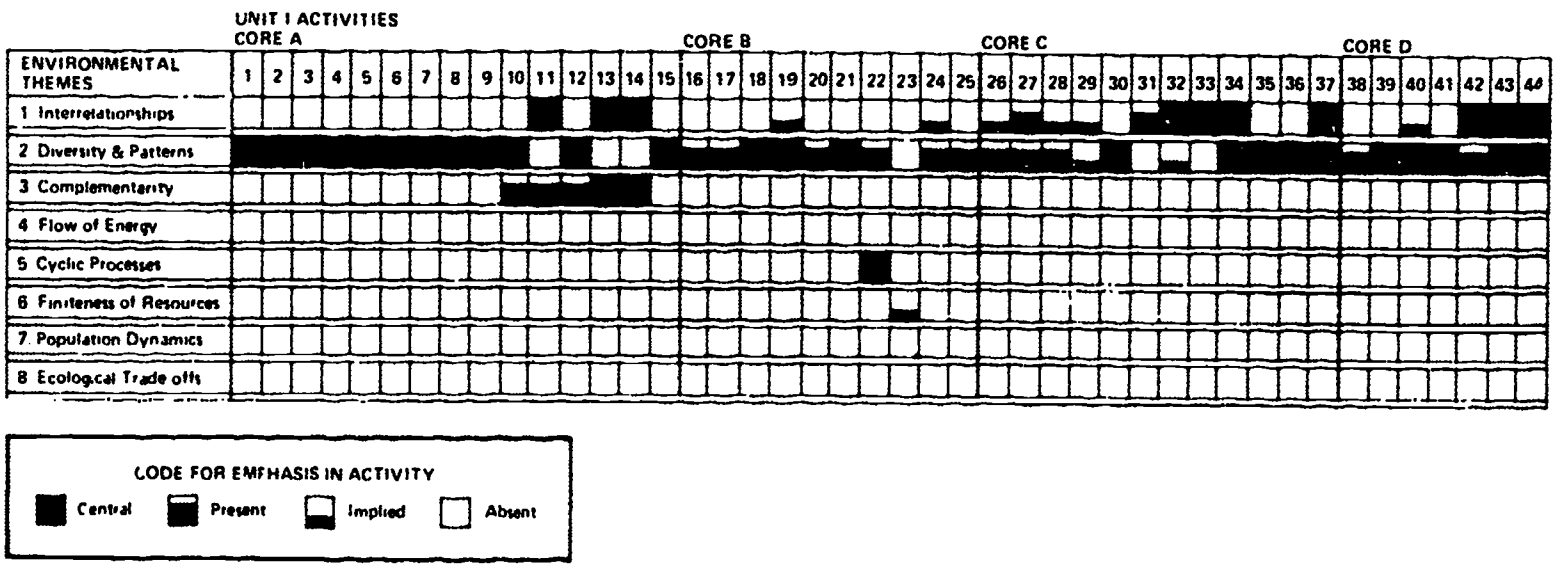
(Continued)

ENVIRONMENTAL THEMES

The matrix shown in Table 1 shows the emphasis given to some of the eight environmental themes intended to be developed in the course of the total sequence of units in ME AND MY ENVIRONMENT. These themes are listed in no particular order of difficulty, but simply represent basic dimensions of ecological understanding. For each of the themes, a several-paragraph description is provided for teachers in the introductory section of the materials. The same information was provided to writers of the activities. The intent of the materials has *not* been to develop the level of generalization and understanding in students implied by the descriptions. Instead, the themes serve to organize for the teacher some of the environmental concerns in the materials. Students may become aware of specific information that illustrates the "Interrelationships of Environmental Components," for example, but they are not expected to be aware of the general category.

In Unit I, three of the themes received major emphasis in more than one activity. The "Diversity and Patterns" theme received some degree of emphasis in all but six of the forty-four activities. In almost half of the activities, aspects of the "Interrelationships" theme received attention. The theme of "Complementarity" received some development in five activities in the first core of Unit I, covering some ten days of instruction. Other units tend to emphasize other themes heavily. For example, in Unit VI, a proposed fourth year of instruction, the themes of "Finiteness of Resources" and "Population Dynamics" will receive their fullest development.

TABLE 1: MATRIX OF EMPHASIS ON ENVIRONMENTAL THEMES



The matrix shown in Table 2 reveals the amount of emphasis placed on various inquiry skills. These skills, along with the problem-solving skills, represent a central part of the program, equal in importance to the subject-matter content. Because of their centrality, some explanation of these skills beyond simple definitions is in order. A hypothesis accepted by the project staff is that the first eight inquiry skills are ordered by level of difficulty and complexity, from observing at the lowest level to applying at the highest level. Furthermore, it is presumed that the skills are hierarchical--each higher skill involving the use of lower skills in the sequence. Of course for any specific skill, such as observing, tasks could range from very simple and easy ones to quite complex, difficult, and demanding ones. It is assumed that the EMH child will profit from a design that proceeds, where possible, from simple to complex, from concrete to abstract, and from the familiar to the unfamiliar. It is also presumed, however, that a child might be expected to perform a higher skill without necessarily being able to do a lower one in specific instances; for example, to describe something observed without necessarily identifying it.

The actual relationship of these skills is unknown, and it is also unknown how developmental in nature they might be. The implication of the label "developmental" is that the skills are not readily acquired by direct instruction at the time chosen by the school. They depend, instead, on the cognitive development of the child.

Several additional skills are listed with the eight inquiry skills, but are not presumed to fit at a particular place in the hierarchical relationship. These skills--guessing, speculating, predicting, divergent production, and value judging--are of a somewhat different order than the preceding abilities.

As can be seen in Table 2, one or more of these skills received *central* emphasis in thirty-eight of the forty-four activities in Unit I. In only one activity in the year of instruction is there no emphasis on any of these skills.

A summary of the number of times each of the first eight inquiry skills is clearly present or central in an activity reveals the following pattern of emphasis:

	Present	Central	Total Times Emphasized
1. Observing	6	24	30
2. Identifying	7	16	23
3. Associating	6	18	24
4. Describing	7	14	21
5. Comparing	13	14	27
6. Translating	1	3	4
7. Inferring	5	4	9
8. Applying	1	7	8

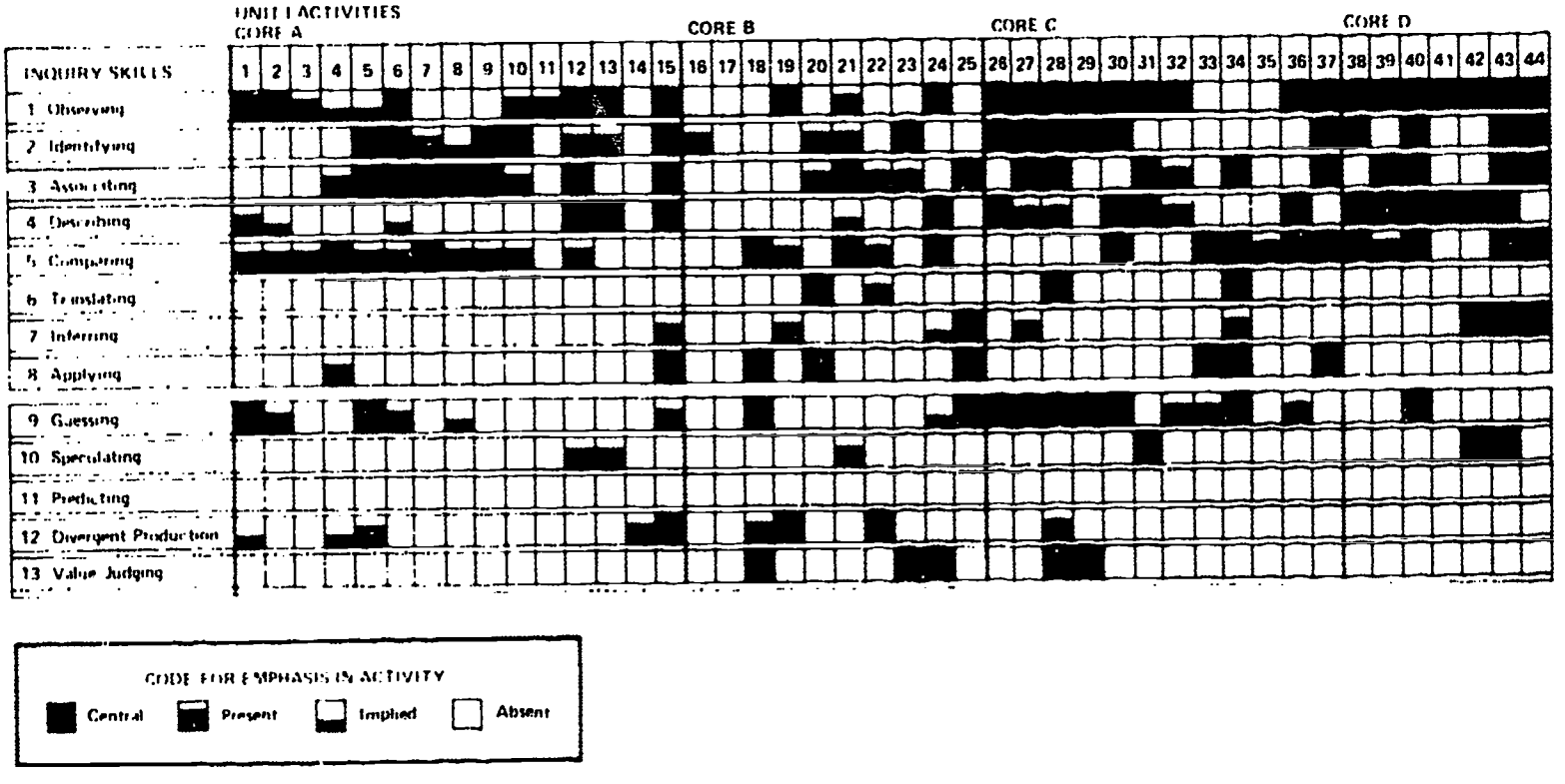
It can be seen that the first five skills received more emphasis than the other three in this unit; of the first five, those skills felt to be basic and requisite to success in developing more complex skills received the most emphasis. Without exception, when a high-level skill (5-8) received central emphasis, a lower skill (1-4) also was centrally emphasized to ensure a degree of success for each student in the activity. The intent was to introduce a few of the skills at a time and gradually to add skills that were more complex. As the matrix shows, central emphasis in Core A of the activities was almost entirely on the first three skills. By the end of Core B, each of the skills had received central emphasis in some activity. Core C reemphasized the basic skills and Core D provided an expanded emphasis on groups of skills.

The three inquiry skills felt to be the most difficult and complex (translating, inferring, and applying) received least attention in the first year of ME AND MY ENVIRONMENT. They will be more fully developed and emphasized in the second and third years. The intent, however, is to redevelop the whole range of skills in each year of instruction, though more rapidly each year that they are repeated.

Guessing and speculating occur in twenty-five activities and are stressed in fourteen of these. Predicting is not emphasized in Unit I. It is presumed that lack of background will require most students to guess rather than to speculate or predict. These activities are related to value judging in the

(Continued)

TABLE 2: MATRIX OF EMPHASIS ON INQUIRY SKILLS

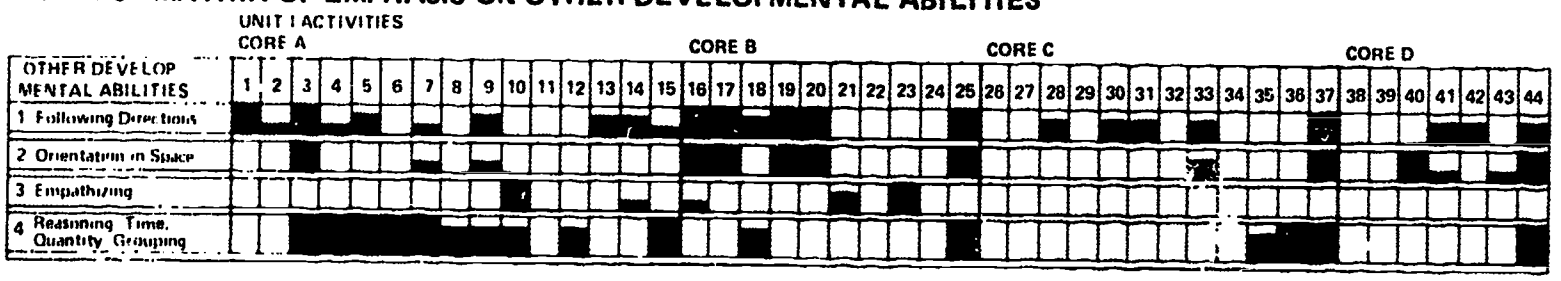


sense of eliciting and accepting individual points of view. They are also quite useful in providing the teacher some insight into the thinking of the child.

Opportunities for divergent production (creating, generating many ideas or categories) occur in three of the four cores of Unit I. Of the nine activities emphasizing this process, however, it is a central feature of only three activities. In further units of ME AND MY ENVIRONMENT, efforts will be made to emphasize this skill more frequently. The application of this process comes in generating or considering alternatives, and thus it is quite important. The use of divergent production is implied in many of the inquiry skills, yet most activities involving these skills compile the responses of the group rather than encourage individual students to apply it.

The opportunity for individuals to apply their own sets of values to a situation occurs only five times in the year. While values are assigned in other activities, students are directed to a particular conclusion. A related aspect of valuing, however, occurs frequently in the unit. Often the teacher is advised to accept any answer the student can justify. This provides an acceptance and endorsement of the student's point of view similar to the acceptance of values implied by this category.

TABLE 3: MATRIX OF EMPHASIS ON OTHER DEVELOPMENTAL ABILITIES



OTHER DEVELOPMENTAL ABILITIES

During the time the matrix of skills was first developed and applied, it was acknowledged to be incomplete. Several additional skills have since been identified as important, and a planned sequence of emphasis has been programmed into the materials. In the final revision of Unit I, for example, the functional abilities of following directions, orientation in space, and quantitative and categorical reasoning were emphasized. Some of these skills are close in kind to the abilities Piaget has addressed in his studies. The intent in the curriculum is to allow opportunity for involvement in tasks calling for these skills. There is no expectation that students will acquire them from direct instruction.

Another skill that has been identified as important—empathizing—relates to the second and third year of instruction. It received little emphasis in Unit I, for it was felt children of thirteen or fourteen are likely to be highly egocentric.

Still other skills will probably be identified as the curriculum evolves. One that has not yet been formally defined, for example, is the ability to deal with an increasing number of variables. Thus far, writers and staff have simply been concerned to avoid activities that require dealing with more than two or three variables at a time. Whenever possible, tasks are broken down so that one variable at a time is considered.

As Table 3 shows, following directions is a skill that is emphasized (but not the subject of direct instruction) in half of the activities in Unit I. It receives special stress in eight of the forty-four activities. Not only is emphasis distributed in every core of the materials, it is placed on a variety of directions: written and verbal, single and multiple, and with written and physical responses required.

Fourteen activities involve ability to orient oneself in space, either in relation to compass directions, or in relation to left and right. In nine of these activities, special attention is directed toward acquiring this ability. While both this skill and following directions are more heavily emphasized in the first half of the unit, emphasis is repeated at intervals throughout.

Opportunities for empathizing occur in six activities in Unit I. In three activities, this skill represents a central focus. As stated earlier, the frequency of emphasis on empathizing was kept intentionally low in Unit I because of what is believed to be the egocentric perspective of students at ages thirteen or fourteen.

Sixteen activities attend to the reasoning dimensions of time, quantity, and grouping. One-fourth of the activities in the unit deal directly with these skills in the attempt to provide many concrete experiences that may enhance their development.

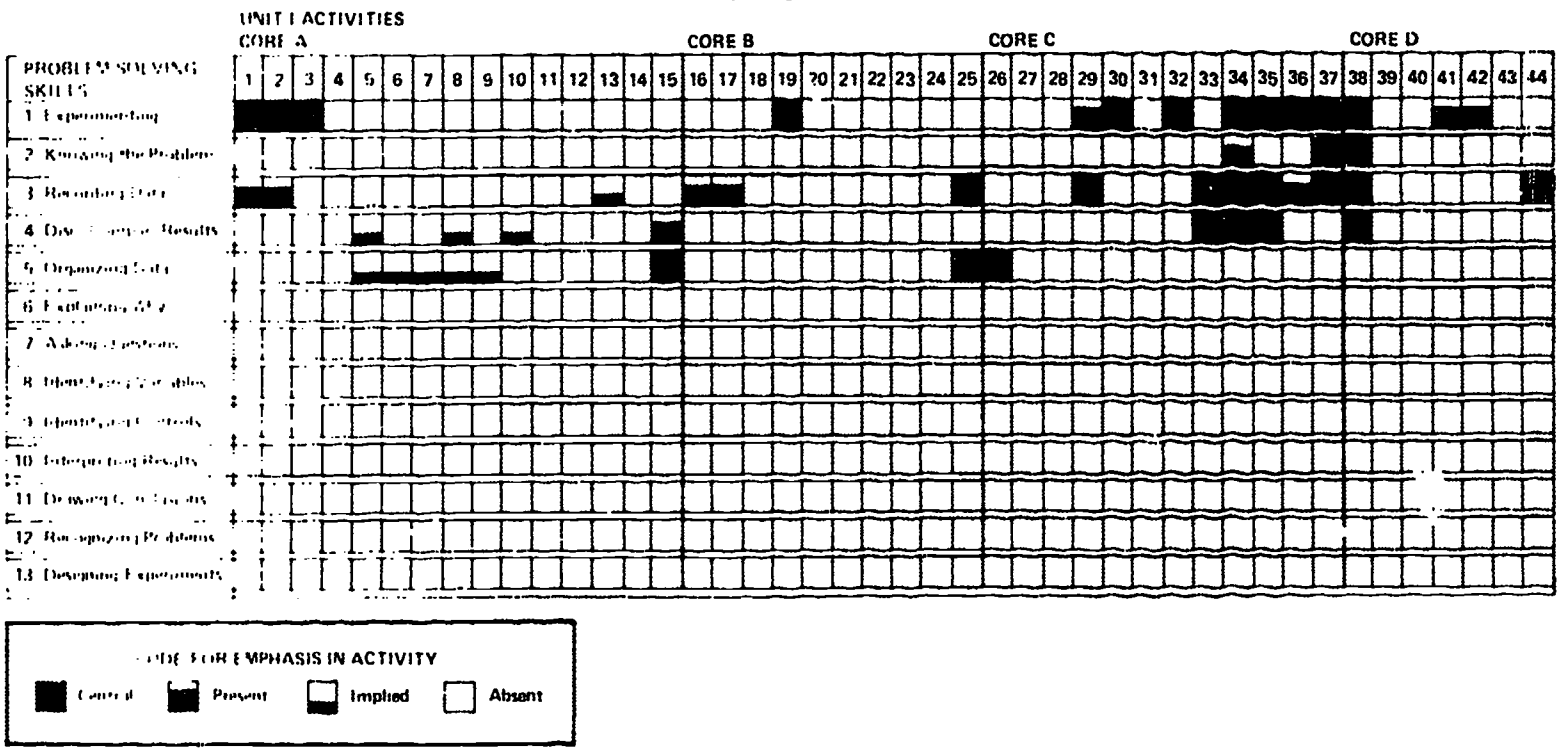
PROBLEM-SOLVING SKILLS

The matrix shown in Table 4 reveals the amount of emphasis placed on various problem-solving skills in Unit I. By intent, this complex set of skills did not receive as much emphasis or development in Unit I as it will in subsequent units. It was intended originally to emphasize many of these skills in the first year of instruction. Unit II, which was originally planned to be taught in the spring of the first year, develops many of these skills. The expansion of Unit I to develop the many facets of the program more adequately resulted in a full year's work. Few of the Unit I activities take on the character of a science experiment in the more formal sense, which is to be the vehicle used to develop the skills of problem-solving in this program.

Opportunities for experimenting in the sense of trying things out to see what will happen occur fourteen times—in about one-third of the activities in the unit. In most of these activities, and in a total of fourteen of the forty-four activities, students record data and refer to it in some way.

On four out of eight occasions for treating group data, a central emphasis of the activity is to encourage students to compare the results they recorded or combine them in some way. In seven cases, opportunities are stressed or implied by the activity to explore students' reasons for interpreting outcomes in different ways. In this unit, the skill is dealt with informally rather than to seek justifications in a systematic way from all students. *(Continued)*

TABLE 4: MATRIX OF EMPHASIS ON PROBLEM-SOLVING SKILLS



VARIETY OF STUDENT INVOLVEMENT

The matrix shown in Table 5 deals with structuring variety in the ways students are involved in activities. Because these categories are self explanatory, no definitions are provided for them. One-fourth of the activities suggest or require students to get out of the classroom and usually to observe or collect things out-of-doors. These activities occur at intervals of four to seven activities throughout the year.

Over half of the activities have as a central feature the use and manipulation of real objects, with the heaviest emphasis occurring in Cores A and C. In only two cases do as many as four activities go by without this kind of involvement. The introduction into the classroom of pets, plants, and a pond early in the year further ensures the continuous emphasis on physical involvement with living things.

Students use worksheets in seventeen activities, half of them to provide the teacher with a means of assessing student background information or understanding. At nine points in the year, detailed assessment activities occur, providing a more formal and systematic check on student development and learning.

Teacher demonstration, which occurs at seven widely separated points in the unit, typically is accompanied by student performance or participation in the demonstration.

Audiovisual aids used in the unit include a poster, a study picture set of 32 plant and animal pictures, a card game of 72 pictures, a model, 53 slides, 3 filmstrips totaling 191 pictures, and 2 sound tapes comprising 45 minutes of recorded sounds, as well as a Polaroid camera for student use, and a microscope.

Another aspect of involvement relates to the size of the groups students work in. In two out of three activities throughout the unit, students spend at least part of the period in full class discussion. Also, in two out of three activities, a central feature is that students work individually or in teams, actively doing something with their hands. Out of the thirty-two activities that involve subgrouping, in seventeen of them students work individually while in fifteen the students work in teams of four or fewer.

To develop the sixty-three dimensions reflected in these five matrices in a curriculum, and to do it in a consistent fashion, is an incremental task. It cannot be done all at once. This effort has proved to be one of the more difficult aspects of curriculum construction. Nevertheless, the matrix analysis of one unit that has been released for general use suggests that the program does systematically emphasize a complex web of ideas, skills, and activities. ■

TABLE 5: MATRIX OF EMPHASIS ON VARIETY OF STUDENT INVOLVEMENT

UNIT I ACTIVITIES

VARIETY OF STUDENT INVOLVEMENT	CORE A															CORE B										CORE C										CORE D									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
1. Out of Classroom				■	■						■		■	■					■	■								■	■					■						■					
2 Use of Real Objects	■	■	■	■	■	■					■		■	■				■	■			■					■	■		■		■	■	■	■	■		■		■	■		■	■
3 Use of Worksheets	■	■	■	■			■								■	■	■		■	■								■		■		■		■			■				■			■
4. Assessments			■				■								■		■	■							■						■			■			■							■
5 Use of Demonstration						■			■									■										■	■				■									■			
6 Audiovisual Aids	C	C	C	S	C	C	G	P	G	P		P	P		SPC				S		F		S	S	S	F	T	T			C		M			C	T	S			FC			SC
7 Discussion				■	■	■	■	■	■	■	■	■	■	■	■			■	■	■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	■	■
TEAM OR INDIV WORK																																												
8 Experimenting	T	T	T									I							I				T								I	I	T	IT	T		I	I		T	T		
9 Manipulating	T	T	T				I		T		I		I	T									T				T		I		I					I				I	T		
10. Constructing				I			I				T			T								I	T									I											
11. Collecting					T								T	T													T	I					T					I					
12 Measuring																I	TI	I							I			I				I	IT	I	T	I								
13 Gaming							I	I	T	T										I												I						I						

CODE FOR EMPHASIS IN ACTIVITY·

EMPHASIS	A/V AIDS	GROUPING PATTERN
■ Central	C Chart	T Team Effort
▨ Present	P Pictures	
▭ Implied	S Slides	
□ Absent	F Filmstrip	I Individual Effort
	M Model	
	G Game	
	T Tape	

revisions–kinds and degrees

THE EVOLUTION OF A UNIT OF INSTRUCTION

[illegible] examples of revisions that occur over [illegible] development, examples are drawn from the [illegible] ME AND MY ENVIRONMENT.

Changes in Unit I did indeed occur over two field tests and [illegible] twenty-six activities to forty-[illegible] original activities survived in relatively [illegible] commercial edition. Table 1 charts the [illegible] deletion, revision, and reorgani[illegible] three versions of the materials. The [illegible] activities in the commercial [illegible] were first inserted into the program, and [illegible] that occurred at each revision. For [illegible], 1.22, in the middle of the commercial [illegible] comparing activity (I-26) in version [illegible], (I-1) in the original materials [illegible], 1.23 in the third version was the last [illegible] materials. The sequencing of [illegible] modification.

[illegible] reorganization reflects major conceptual [illegible]. The first version of the materials listed five [illegible], along with seven of the ten associated [illegible], from the start. This objective, [illegible] relationships, involved [illegible] about distance, mapping and [illegible] scale. These understandings [illegible] the main thrust of the program, [illegible] the time necessary to develop them fully.

BEST COPY AVAILABLE

Two of the five original objectives remain as unit goals in the final version. This version gives six unit goals and eighteen core objectives that are distributed across the four clusters of related activities in the unit. The other two original objectives can be found stated in a slightly different way as the first four objectives for Core A of the final materials. An examination of the objectives for Unit I indicates a refinement of them toward more clarity and specificity, as well as a considerable expansion of scope and purpose.

The expansion of Unit I is shown by the fact that thirty-one new activities were conceived after the first twenty-six were written and tried out. Combining and dropping activities took a toll of thirteen. Revisions typically required the expansion and elaboration of activities. Thus the unit grew

from 26 activities in 125 printed pages,
to 29 activities in 219 printed pages,
to 44 activities in 568 printed pages.

The most extensive revision was made in producing the second version. Eleven activities were dropped from the unit and about nine others were considerably changed. In the second revision leading to the commercial edition, two more activities were dropped and about six were substantially changed.

These changes were based heavily upon feedback from field trials. The examples included in this article illustrate how the information was organized for review. The process of compiling the data is described in Part III of this report.

(Continued)

ME AND MY ENVIRONMENT
1974 COMMERCIAL EDITION OF UNIT I COMPARED TO EARLIER EXPERIMENTAL VERSIONS

Commercial Version of Unit I Core and Activities Titles and Sequence	Activity Number	Final (3rd) Version	Revised (2nd) Version	Original (1st) Version
CORE A: INVESTIGATING THE VISIBLE ENVIRONMENT				
The Cube Float Puzzle	1·1	New activity	.....	.....
To Fizz Or Not To Fizz	1·2	New activity	.....	.
Clues To Success. Background	1·3	New activity	... ..	. ...
Introduction To Sorting	1·4	Revised from	1·13 (new)	. . .
Sorting Things In Our Environment	1·5	Revised from	1·14 revised from	1·6
Categorizing In Terms Of Living/Nonliving	1·6	Revised from	1·15 revised from	1·9
Clues To Success: Understanding And Background	1·7	Much revised from	1·22 (new)	.. .
Defining The Word Environment	1·8	Much revised from	1·3 revised from	1·3
Environmental Rummy	1·9	Much revised from	1·16 revised from	1·8
Some Animals...	1·10	New activity	.....	.....
An Animal In Class	1·11	Revised from	1·1 revised from	1·20 and 1 21
Some Plants...	1·12	New activity	.....	....
Plants In Class	1·13	New activity	.....	.....
A Pond In The Classroom	1·14	Revised from	1·2 (new)	.. ...
Clues To Success: Understanding	1·15	New activity	.....	.....
CORE B: LANDMARKS IN THE VISIBLE ENVIRONMENT				
The Grid Game	1·16	New activity	.....	.....
Clues To Succes Background	1·17	New activity	.....	.....
Measurement In My Environment	1·18	Revised from	1·21 much revised from	1·11
Establishing Environmental Landmarks	1·19	Revised from	1·23 (new)	.....
Landmarks In My Environment	1·20	Revised from	1·24 revised from	1·16
Some Houses...	1·21	Revised from	1·25 (new)	. ....
A Recycled Spaceship	1·22	Minor change from	1·26 revised from	1·1
To The Moon In Our Room	1·23	Minor change from	1·27 minor change from	1·26
Far Out Environments	1·24	Minor change from	1·28 much change from	1·2
Clues To Success: Understanding	1·25	Much revised from	1·29 (new)	.....
CORE C: SENSING THE INVISIBLE ENVIRONMENT				
Some Sounds...	1·26	Revised from	1·6 revised from	1·4
Environmental Orchestra	1·27	Revised from	1·7 (new)	.....
Sounds Around Us	1·28	Revised from	1·8 (new)	. ....
Sniffing Around	1·29	Revised from	1·9 much revised from	1·5
A Strange Feeling	1·30	Revised from	1·10 (new)	.....
Clues To Success: Background	1·31	Revised from	1·11 (new)	.....
Temperatures Affect All Living Things	1·32	Revised from	1·18 (new)	.....
Reading A Thermometer	1·33	Revised from	1·19 (new)	.....
Temperatures In My Outside Environment	1·34	Revised from	1·20 revised from	1·10
Weighing With A Balance	1·35	New activity	.....	.....
Which Is More?	1·36	New activity	.....	.....
Clues To Success: Understanding	1·37	New activity	.....	.....
CORE D: LOOKING AT THE INVISIBLE ENVIRONMENT				
Look Quick	1 38	New activity	.....	.....
Taking A Closer Look	1·39	Revised from	1·3, part 2 (new)	.....
Zoom In..., Zoom Out	1·40	Revised from	1·5 (new)	.....
Entering The World Of The Invisible	1·41	New activity	.....	.....
Life In The Invisible World	1·42	New activity	.....	.....
What Is Environment?	1·43	Revised from	1·12 (new)	.....
Clues To Success: Understanding	1·44	New activity	.....	.....
Summary of changes reflected in the table:		44 activities in 568 pages added: 16 activities	29 activities in 219 pages dropped: 2 activities added: 15 activities sequence: extensively modified	26 activities in 125 pages dropped: 11 activities sequence totally changed

THE NATURE OF CHANGES

The extensive nature of the changes in Unit I was not a matter of caprice, nor were these merely cosmetic changes that could have been made without the immense effort of field-testing and data collection. Documenting the basis for specific changes rarely occurs, however, and for good reason. Expert judgment is a large component in many of these decisions, and an explanation would require considerable background information. Attention is fixed on the program and how to improve it; therefore, a careful record of the meetings and decisions of project staff is not usually kept. Some of the decisions for revision could be reconstructed only by repeating the developmental process and reviewing the data for a second time. Thus, the "whys" of revision typically remain a hidden dimension of curriculum development.

Some examples of changes and their justifications are provided in the remainder of this article simply to illustrate the kinds and degrees of change that have occurred in the development of ME AND MY ENVIRONMENT.

At first glance, it would seem possible to review every activity and categorize it by the amount of change that took place. The attempt to apply a set of categories (such as none, minor, significant, or extensive change) soon leads to abandoning the effort. Sometimes seemingly minor changes made the difference between an activity that succeeded and one that failed. Some activities were considerably expanded into a series of small steps deemed necessary in order for all students to grasp the idea. However, the substance of these activities often remained essentially the same. Almost every activity was modified to include a section at the beginning and the end explicitly relating the activity to others in the series, and clarifying its purpose. This effort to communicate the "why intentionality" of activities to students and teachers was extensive.

EXAMPLES OF CHANGES IN ACTIVITIES

Activity 22 (A Recycled Spaceship) illustrates a major conceptual change. The original activity opened the unit. It developed the concept of a model, introduced the world globe as a model, and attempted within a broad definition of environment to help the student see where he was in the context of the whole world. The activity led toward looking at a number of habitats.

In the second version of Unit I, the activity became a culminating one where the students discussed many of the habitats they had looked at earlier and looked for likenesses between the earth and the spaceship. The "model" concept was dropped entirely and the concept of "recycling" was developed. In the final version the activity culminates the first half-year of instruction. The concept of "recycling" is developed more fully. The concept of "habitats" has been dropped and reference is made to the ways people live in different places around the world. Other than these changes, the activity is much like the second version. Unlike the deletion of the concept of a "model" from the unit, the omission of the term "habitat" represents an effort to simplify rather than to exclude. While the more technical aspects of the term "habitats," as well as the term itself, were deleted, the concept is introduced by referring to different environments in different places for different creatures.

Activity 29 (Sniffing Around) illustrates several kinds of changes. The activity seeks to establish smells as an important component of the environment. In the original version the students received a worksheet picturing a dozen scenes, including a bakery and a wharf. Students were to identify familiar scenes and think about how each place smells. This strategy uses one sense modality (sight) to try to address another (smell). It is abstract and dependent upon the range of experience of students.

Some teachers indicated that the worksheet was confusing to students; many teachers suggested having bottled smells. Content reviewers also pointed out that the purpose and conclusion of the activity were trivial, inasmuch as students already know that they perceive the environment by smelling. Consequently, the pictures were dropped and the activity was rewritten, using containers of a wide range of bottled odors.

(Continued)

Figure 1 provides a sample of the detailed direction to writers for the first revision of Unit I. It indicates the way the conclusions drawn from feedback were relayed from project staff to writers. Writers also had access to a variety of raw data.

The bulky file on the first field test of this activity contained:

- –specific criticisms from reviewers,
- –detailed notes of observers in four classrooms,
- –responses of all teachers to a 20-part questionnaire specific to the activity,
- –summaries of teacher responses to each questionnaire item,
- –news clippings that refer to smells and their effect on people,
- –annotated copies of the activity from teachers' manuals,
- –samples of student responses on worksheets from each class, and
- –student responses to a question about their interests in school and science.

FIGURE 1. Excerpt from Detailed Directions to Writers for First Revision of Unit I

ACTIVITY 9. "SNIFFING AROUND" (Subsequently 1-29)

Drastic revision of present Activity 5. As reviewers point out, students already know that they perceive the environment by smelling, thus the purpose and conclusion are trivial. The activity should cultivate and encourage this idea, but help students realize that different environments smell different for various reasons, that environments can be (sometimes) characterized by how they smell, that smells can give valuable clues to the characteristics of environments, etc. Value judging and debate about pleasant vs. unpleasant smells should be developed. The worksheet of pictures should be dropped! There are all sorts of neat smells that could be bottled for identification by students. They could be asked to speculate on what the environment might look like that the smell came from. They could be asked to collect some smells for others to identify. A field trip to smell the environment could be included—teacher suggestions should include crushing plants, etc. Be sure to name and label plants and other sources of smells. Students could be given a task: go out and find a smell that you never smelled before—and bring it back for the rest of us to smell. Much of our aesthetic appreciation is based on how a place smells. Can we take the students to a wilderness, fisherman's wharf, a sawmill, etc., in bottles? How about having them sand some wooden blocks to see if all kinds of wood smell the same? We might even throw in an information film loop showing a bunch of organisms that rely mostly on smell for navigation, mate-finding, etc. Do these ideas stink?

Figure 2 illustrates the variety of details and recommendations that resulted when the staff reviewed feedback from the field test of the second version of this activity. As can be noted, a worksheet was developed to use in conjunction with the odor containers. This worksheet had the following format:

Smell	Like it	New to me	Name of smell
Jar 1	Yes No	Yes No	

The second field test revealed that this worksheet was difficult for students to complete. Specifically, students tended to circle "no" if the smell was new to them, whereas the question required a "yes" answer. Spelling the name of the smell completely frustated many students. A new format was therefore developed for the worksheet, and instructional strategies were devised to get the names of possible smells on the board for students to copy.

The activity content was changed in the final revision, too. The concept of "environment" needed more direct development. More reference was needed to the smells of spoiled food and to the smells of household materials that could be dangerous. The activity required division into parts so that teachers would spend several days on it. Teachers had to be urged to take their students outdoors to collect more smells to share. The idea that smells can warn of danger needed more emphasis. Additions were made to the activity to incorporate all of these suggestions for improvement.

Activities 32-34 illustrate several other changes made in response to feedback. In the original version of Unit I, one eight-page activity dealt with temperature. Because much more was found necessary in developing the concepts involved, the second version was expanded to three activities comprising twenty-four pages. In the third version, the three activities comprise thirty-five pages. This fourfold expansion parallels the overall increase in size from the first to the third version.

Included in this considerable expansion of instruction related to temperature were the following changes:

An entire activity developed the idea that temperature affects living things. Two experiments were included in this activity to emphasize the development of inquiry and problem-solving skills. The concept of differential absorption of heat by light and dark colors was deleted from the activity.

A second activity called for students to construct thermometers and practice reading them. Students were also given practice with real thermometers and the chance to compare their readings with others.

An assessment of the students' ability to read the thermometer accurately precedes the collection and recording of indoor and outdoor temperatures in a third activity. The graphing of daily temperatures over a period of time and a comparison of these records was added to the activity. A list of specific steps and cautions to use in getting accurate readings was also included.

Further details on the kinds of changes that occurred in revision and the evidence that contributed to the decisions are reflected in some of the illustrations used in Part III of this report.

FIGURE 2: Staff Summary of Feedback from Field Trials of Second Version of Activity 1-9: Sniffing Around

GENERAL SUMMARY AND RECOMMENDATIONS:

1. This activity proved to be highly successful in interesting students. The worksheet was difficult and frustrating to fill out, however, and should be revised.
2. More strategy is needed to develop familiarity with household materials that could be dangerous, and recognition of the smell of spoiled foods.
3. The term (and concept) "environment" needs to be further developed in this activity.
4. Few used the optional activity because of the weather. It needs to be written so that teachers understand that it could be done indoors at school.
5. The activity should be expanded and divided to indicate clearly that it takes at least two days of class. (Most spent one day, and less than eighty minutes.)
6. Preparation time was extensive: one to three hours for two-thirds of the teachers. All said it was worth it, but the planning guide and page 70 should indicate time needed.

SPECIFIC DETAILS

1. Worksheet (and accompanying slide) revisions:
   a. First column: Change "jar" to "carton" or omit it entirely. The strategy used the term "carton" only.
   b. Issue: Is it wise to have three things to do on the worksheet?
      Con: students have difficulty attending to more than one or two things at a time.
      Pro: perhaps experiences should be planned requiring attention to more than two things, if kept simple.
   c. Idea for second column: Instead of yes-no, use Mmm and Yuk, with drawings of faces (using the new symbol for poison for the bad smells).
   d. Column 3 (new to me): The heading for this column should be changed from "new to me" to "smelled before." An observer noted that if students hadn't smelled something before they tended to circle "no" whereas the question required a "yes" answer, thus causing confusion. (This was also noted by teachers 39 and 49.)
   e. Spelling is apparently a large and frustrating problem. Putting the names of smells on the board is one solution. Another idea would be a worksheet that lists the names of fruits, vegetables, other foods and seasonings, kitchen supplies, bathroom supplies, and dangerous materials. These could be sorted into categories as an additional strategy. Lines could be left for adding other things. On a second day, and using the worksheet, students could go around the stations again and check themselves by finding the name that has been placed on each station. That way they can sniff again and learn the unfamiliar smells and complete their worksheets.
2. Summary of specific questions on feedback sheet:
   "Was recording on Worksheet 1-2 too difficult?" 12 yes, 8 no. Many said students couldn't spell or write in name of smell. Column 3 (new to me) confused them.
   "Did many say some smells were new?" 12 yes, 6 no. Smells mentioned most often as new were oil of cloves, sage, Clorox, and ammoni Another smell idea would be to use oil, perhaps fuel oil. Additional smells used other that those listed in Guide were: cinnamon, vinegar, vanilla, cloves, crayons, mustard, pickle, cleaning fluid, Vicks, coffee grounds, laundry soap, rotten egg, toilet bowl cleaner, Lysol, candy, deodorant, rum extract, peppermint, hydrogen peroxide, orange marmalade, Noxema, Ben Gay, pepper, piece of Christmas tree (fir), moldy canned peas, turnip.
   "Did you do the optional activity?" 5 yes, 15 no. Eight said the weather was too cold or snowy.
3. Hidden smell (page 70) incense worked best. Oil of wintergreen or peppermint elicited no comment in several classrooms. Recommend they be deleted from suggested smells.
4. Addition to strategy (bottom of page 73): Refer to smell of ammonia and tell students that whenever they are smelling a strange material, it would be safer to smell the cap or just fan the top of the container rather than stick their noses in it. Demonstrate this.
5. Strategy extension:
   There are only three questions on the point of the lesson (page 73 bottom and top of page 74). Extend lesson to develop ideas of warning signs, poisons, and spoiled food.
6. Evidence of strategy weakness:
   Page 74, "Why are smells an important part of our environment?" "I had to ask a number of leading questions to get any response." (38) A specific section should more graphically illustrate that smells warn us of imminent dangers.
7. Revision ideas:
   a. Start with fewer smells on the first worksheet and have them all work together so they learn how to use the worksheet. Then do more smells on a second day.
   b. Include second worksheet (or extend activity) dealing with names of household materials that are dangerous. 29 reports that his children were *unable* to read the names such as *ammonia, alcohol,* and *Clorox,* etc.
   c. Perhaps a sequence could be added about dangers you can't smell—such as carbon monoxide from car, heater, or using charcoal indoors.
   d. Perhaps several foods could be used in both fresh and stale or spoiled form so that students could learn to recognize how smells change.
8. Clues to Success suggestions:
   Teachers should wear the same perfume or aftershave for several weeks before this activity--then use it as a "smell" and see how many identify it with them. See anecdote for teacher 39.
9. Organization of activity:
   a. Eight or ten teachers commented on the procedure for smelling odors and completing the worksheet. Most set up stations and felt this worked fine. The most satisfied seemed to have as many (or more) smells as students, and several who didn't pointed out this was crucial for success.
   b. The spelling problem interfered with the approach of having each student on a different smell. For this reason teacher 33 used the following strategy, which might be a good way to introduce the worksheet although it isn't a good plan for the whole exercise.
      "I passed around 1 carton at a time. Each child sniffed the same carton. The slide was projected and we answered 'like it' and 'new to me'. Then I asked each one to say what it was and wrote each new answer on the board. They could then copy the word they thought was correct. Using this method each got to sample 5 or 6 cartons instead of 12, but they weren't frustrated by being unable to spell."

ANECDOTES

Teacher 31: *"This lesson proved so successful in developing language as well as scientific skills that another teacher came in and 'swiped' the whole lesson (cartons and all) to use with her disturbed class today."*

Teacher 36: *"Some didn't know names of smells but could tell me what they were used for. For instance, one boy said sage was the stuff put in turkey dressing. Another said cloves was stuff used in 'punkin pies.'"*

Teacher 39: *"The last carton included some perfume that I wear to school often and an unbelievable number recognized it."* ■

the collection of data

The first two reports in the formative evaluation of ME AND MY ENVIRONMENT have discussed the multiple purposes and uses of data, and the different audiences for various studies. The major purpose of the evaluation effort has been to gather data that directly informs the revision effort of what happened with the materials in use, and of what the conditions of use were.

A second focus of study has been on the abilities and background of the target population and the performance of students on tasks related to instruction. Typically, these studies are broader in scope and longer in range than the first studies, but they also provide a general framework that can be used for the design of the materials. In both cases the developers of the materials are themselves the primary audience for the studies, with the funding agency and the field of education important second audiences.

A third purpose of the evaluation effort has been to gather evidence regarding the overall design and impact of the program. This report records this procedure. Again, the individual studies have been planned to give direction to the revision effort as well as to characterize the student population. Activities described in the first two parts of this report reflect that double utility. This third purpose, which is long range and will not be fully realized before the final report in this series, has as its primary audience the field of education and the funding agency.

A summative evaluation of the materials after they are released for commercial distribution should be conducted by an outside agency to study the effectiveness of the total three-year program. The formative evaluation effort is clearly *not* designed to *prove* the effects of the program under development. Its central purpose is to *improve* the effectiveness of the final product. This distinction needs to be repeated to prevent misinterpretation of the results contained in these interim reports. These results and conclusions relate to the experimental materials and the outcomes of their first and second field trials.

The final product will incorporate, insofar as possible, the corrections and revisions suggested by these findings. Ultimately the standards and criteria these reports contain should be applied to the final product, and judgments of its worth and effectiveness should be reported to users. It would seem appropriate for the publisher and funding agency to initiate jointly such a consumer-oriented study, to be conducted by a disinterested third party.

CONCERNS IN DATA COLLECTION

A primary concern in data collection has been to try out the program in more than one class for each of a number of conditions of use: age of students, level of ability, ethnic and cultural setting, geographic location, and school size and location related to population density. In sampling these conditions, no attempt was made to construct a representative sample, but rather one that would suggest the results were typical.

Much attention was given to the materials *in use,* rather than assuming that the same treatment occurred in all test classes. Information regarding differences in use was used to qualify the results obtained. The data have not in many cases been additive; simply counting the number of times something occurred is often misleading. It is the nature of classroom transactions and day-to-day outcomes that have best directed the revision effort.

Another central concern in data collection has been exploring the reliability and validity of information. Redundancy of data is one way of verifying credibility. For example, the validity of teacher observations has been questioned in some cases. It has been asked, "Can teachers make this particular judgment in the midst of everything else they are doing?" Observers have been asked to complete the same feedback sheets the teachers completed, not because teachers were considered unreliable, but because two rather different perspectives of the classroom could thus be compared. Occasionally two observers would report on the same classroom, and their notes were also compared. Teachers and others were sometimes asked for the same information in different ways to determine whether a consistent view was held. In all cases, the redundancy, or lack of it, helped developers to judge the reliability of the data.

Questionable data was checked in various ways. The methods used by staff and observers for recording results systematically or for interviewing students were checked, and conflicts in data were noted and explored. Paired observations and staff debriefings were also used as a way of increasing the reliability of the reports. Data from students was verified by means of test-retest studies, interviews compared to performance measures obtained in other ways, interviews to obtain explanations of test-item responses, and comparisons between observed and reported behavior.

A number of checks were used in verifying the accuracy of teacher reports. Responses to checklist items on the feedback forms were reprocessed. If little or no variation in ratings occurred, the credibility of that teacher's reports came into question. Teachers whose ratings were always high or satisfactory were looked at more closely to determine whether their feedback had any utility in critiquing and improving the materials. Some teachers' feedback was dropped from all but cursory consideration because of these reasons, or because of a lack of descriptive and critical comments on the activities.

Other indicators of questionable data were teacher records of student performance where performance was uniformly perfect or where no variance occurred. Worksheets were occasionally rescored and judgments were compared to those of an observer. Errors in the teacher's manual became a means for checking the alertness of the teachers and their ability to understand the intent and make appropriate corrections. Observations and interpretation of feedback enabled the staff to make judgments regarding the teacher's degree of fidelity in implementing the materials the way they were intended. Data from classes was weighted according to these considerations.

PROBLEMS OF DATA COLLECTION

The problems of data collection are legion. A few of them are recited here. First of all, mortality of test classes and attrition rate of students reduced the sample studied. Failure of teachers to keep up with the pace necessary in testing the materials, or to complete the year's work, also reduced the data available.

Looking at the group of fourteen classes that tried out the first version of the curriculum over a three-year period, the following response was obtained:

TABLE 1 Completion of Testing and Return of Data by Field-Test Teachers

	No. of Teachers Starting the Year	No. Who Failed to Supply Data	No. Who Fell Behind & Didn't Finish a Unit	No. of Teachers Finishing the Year	Avg. No. of Responses Received from All Teachers Per Activity
Year 1, Unit I	14	1	0	13	12
Year 1, Unit II	14	1	3	10	8
Year 2, Unit III	14	0	0	14	13
Year 2, Unit IV	12	0	3	9	8
Year 3, Unit V	8	0	3	5	5

A similar pattern of response holds true for the group of twenty-one classes that field-tested the second version of the materials.

In the same three-year period, the attrition rate in student enrollment was high. Over one-third of the students left the classes each year. At the end of the third year, less than 27% of the students enrolled the first year remained in the program. Even within a single year of the program, the attrition rate was high enough to represent a significant problem in data collection. When this is compounded by a lack of data due to student absences and by errors on the part of the teacher, the problem becomes even more serious. In the first year of field-testing the first version of materials, a

majority of the data was collected at three times in the school year. Complete data were obtained on 70% of those students enroiled at the end of the year. In the second year of this field test, data were gathered at more frequent intervals throughout the year. Complete data were obtained on 58% of the students for the first unit of instruction in the year. This percentage dropped to 30% for the last half of the last unit tested that year. The results were similar for the third year.

Another problem of data collection is the sheer mass of data to be processed. In the second year of field-testing, almost 1,700 teacher feedback forms alone were received. The procedures for organizing and processing this data are described in a following article.

A related problem is the accessibility of data. The processing of interview data and the adequate documentation of observations are two cases in point. In each case, reducing these rich sources of data to a meaningful set of summaries and conclusions is a difficult task.

The time of requests for information and the timely receipt of data in order to process it before another round of collection and processing is due are all too familiar problems.

COSTS OF DATA COLLECTION

Typically, one thinks of budget-imposed constraints on the conduct of an adequate evaluation. It is true that a number of compromises must be made, but these are a reality of any evaluation effort. Often one's desires for data exceed one's resources. It will be one purpose of the remaining articles to indicate the various costs, benefits, and trade-offs involved in each type of data collected. The final section of this report attempts to capture some of the reality of the classroom not represented in most collections of data.

Evaluation has both direct and indirect costs. The invention, production, administration, and analysis of a questionnaire represents only one set of costs. Also included should be the number of personnel involved in the inventing, planning, and interpreting. The cost in terms of time needed to respond to the instrument should also be considered, as should the impact upon the attitudes and good will of teachers and students.

Not all the kinds of data collected and studied in the formative evaluation are discussed in this section. The preceding sections, for example, have described many forms of analysis applied to the curriculum content. This section deals with the data collected from field-test classes. Table 2 (page 32) lists the variety and sources of this kind of data. Still other forms of data will be treated in later reports. ■

TABLE 2. DATA COLLECTED FROM FIELD-TEST CLASSES FOR THE FORMATIVE EVALUATION OF ME AND MY ENVIRONMENT

	Data	Source	Population Included	Method of Collection	Frequency and Time of Collection	Place Described &/or Results Reported
Student information						
1	Birthdate	School records	All students	Teacher completed form	Once yearly, fall of each year	See reports 1 and 6
2	Sex	School records	All students	Teacher completed form	Once yearly, fall of each year	See reports 1 and 6
3	Ethnic background	School records	All students	Teacher completed form	Once yearly, fall of each year	See reports 1 and 6
4	WISC total IQ (or equivalent)	School records	All students	Teacher completed form	Once yearly, fall of each year	See reports 1 and 6
5	Photograph release	Parents	All students	Printed form sent home by teacher	Once yearly, fall of each year	See reports 1 and 6
6	General description of each student	Teacher	All students	Typed verbal descriptions	Three times in duration of field test	See reports 2, 4, and 6
7	Reason for placement	Teacher	All students	Mailed rating scale	Once yearly, spring of each year	See reports 2, 4, and 6
8	Attitude towards school	Teacher	All students	Mailed rating scale	Once yearly, spring of each year	See reports 2, 4, and 6
9	Awareness of what is going on around him/her	Teacher	All students	Mailed rating scale	Once yearly, spring of each year	See reports 2, 4, and 6
10	Ability to follow directions	Teacher	All students	Mailed rating scale	Once yearly, spring of each year	See reports 2, 4, and 6
11	Ability to work with a group	Teacher	All students	Mailed rating scale	Once yearly, spring of each year	See reports 2, 4, and 6
12	Ability to attend to a task	Teacher	All students	Mailed rating scale	Once yearly, spring of each year (discontinued)	See reports 2, 4, and 6
13	Ability to work with hands	Teacher	All students	Mailed rating scale	Once yearly, spring of each year	See reports 2, 4, and 6
14	Verbal participation	Teacher	All students	Mailed rating scale	Once yearly, spring of each year (discontinued)	See reports 2, 4, and 6
15	Problem solving ability	Student	All students	Paper and pencil test administered by teacher	Once yearly, spring of each year	See reports 2, 4, and 5
16	Cognitive development	Student	All students	Paper and pencil test administered by teacher	Once yearly, spring of each year	See reports 2, 4, and 5
17	Measures of entering knowledge and ability	Student	All students	Paper and pencil test administered by teacher	Periodically through school year	See reports 2, 4, and 5
18	Measures of performance related to instruction	Student	All students	Paper and pencil test administered by teacher	Periodically through school year	See reports 2, 4, and 5
19	Functioning and results from student worksheets	Student	All students	Return of completed worksheets	Periodically through school year	See reports 2, 4, and 5
20	Interest	Student	All students	Teacher feedback form	After each activity is taught	See report 4
21	Absence record	Student	All students	Teacher feedback form	After each activity is taught	See report 4
22	Lack of success	Student	All students	Teacher feedback form	After each core of activities	See report 4
23.	School and work aspirations	Student	50% Stratified random sample	Interview by staff	One time only during site visits, 1971 72	See report 1
24.	Validity study of Unit I post test	Student	50% Random sample from 4 classrooms	Interview by staff	One time only, spring 1972	See report 1
25	Understandings about the environment and reactions to science class	Student	50 Stratified random sample	Interview by staff	One time only during site visits, 1971 72	See report 5
26	Attitudes toward Unit I activities	Student	All students	Questionnaire administered by teachers	One time only spring 1972	See report 5
27	Attitudes toward Unit IV activities	Student	All students	Questionnaire administered by teachers	One time only, spring 1973	See report 5
28	Attitudes toward Unit V activities	Student	All students	Questionnaire administered by teachers	One time only, spring 1973	See report 5
29	Attitudes toward smoking, drinking, and drugs	Student	All students	Questionnaire administered by teachers	Two times only, fall 1972	See report 5
30.	Ability to observe, describe, and categorize	Student	30 Random sample from 21 classrooms	Interview by staff	One time only during site visits, FT 2, 1972 73	See report 5
31	Ability to recognize science apparatus and its use in experiments	Student	30 Random sample from 14 classrooms	Interview by staff	One time only during site visits, FT 1, 1972 73	See report 5
32	Ability to use Polaroid camera and understanding of purpose and use	Student	25 Random sample from 7 classrooms	Interview by staff	One time only during site visits, FT 1 and 2, spring 1973	See report 5
Classroom transactions						
33	Functioning of materials strategies and media in individual activities	Teachers	All teachers	Activity feedback form	One per activity throughout the year	See report 3
34	Functioning of materials strategies and media in other activities	Observers	4 teachers	Activity feedback form written report per activity staff debriefing	One per activity throughout the year	See report 3
35	General functioning and coordination of groups of activities	Teachers	All teachers	Core reaction sheet	Once per core of activities throughout the year	See report 3
36	General operation of class and fidelity to program intent	Staff	All teachers	Site visits	Once or twice (or more) each year	See reports 2 and 3
37	Ability to follow directions	Teacher	All students	Student record of progress	Periodically through the year	See report 6
38	Spatial orientation	Teacher	All students	Student record of progress	Periodically through the year	See report 6
39	Problem solving	Teacher	All students	Student record of progress	Periodically through the year	See report 6
40	Grouping	Teacher	All students	Student record of progress	Periodically through the year	See report 6
41	Cognitive development	Teacher	All students	Student record of progress	Periodically through the year	See report 6
42	Interest in science	Teacher	All students	Student record of progress	Periodically through the year	See report 6
43	Understanding of concepts	Teacher	All students	Student record of progress	Periodically through the year	See report 6

data from teachers, students, and observers

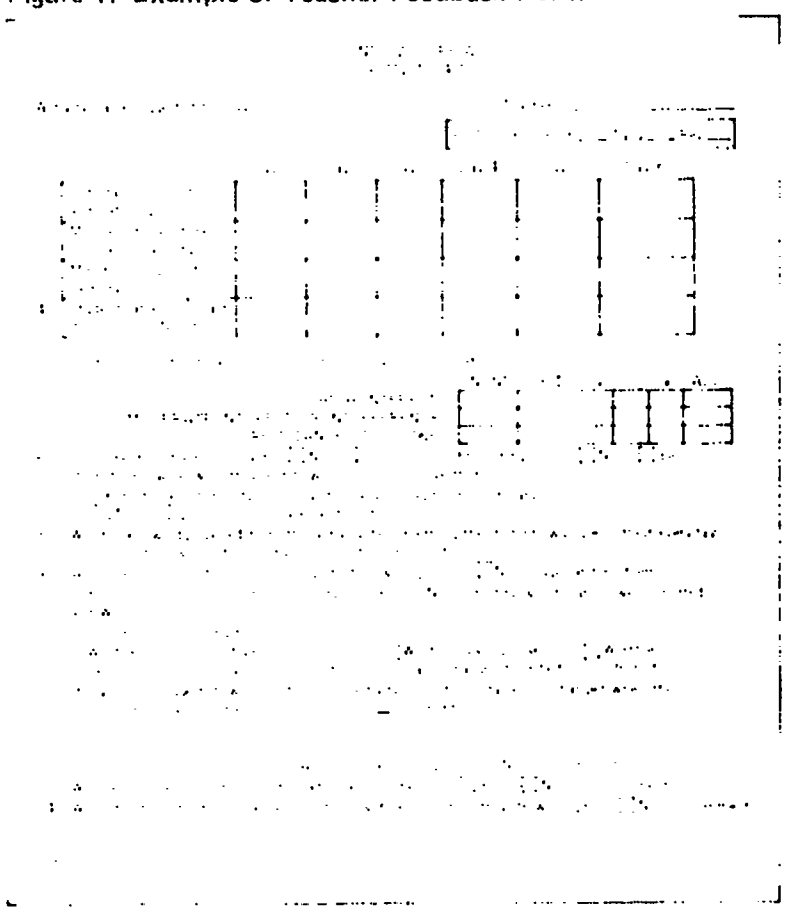
DATA FROM TEACHERS

In the field trials of as totally new and untested a curriculum as the experimental editions of ME AND MY ENVIRONMENT, detailed information is needed about the implementation of each activity. The major source of much of this information is the teacher. To facilitate the teacher reports, a questionnaire was developed that gathers the necessary data efficiently while still eliciting as many original judgments and suggestions as the teacher can supply. This instrument, given the name "feedback form," has undergone several revisions and is shown in its final form in Figures 1 and 5.[1] Each teacher was asked to fill out a feedback form immediately after completing each activity. On the average this was twice a week, involving approximately fifty feedback forms per teacher during the academic year. A total of 217 feedback forms of this type were developed.

Teacher Feedback Form: The feedback form, requesting essentially three kinds of information, was designed almost entirely to serve the purposes of revision writers. The first section asked the teacher for judgments of the overall value of the activity. Questions regarding the date the activity was taught and the time spent were included to determine the average number of days and minutes spent on each activity. The grid for these questions worked very well.

[1] *None of the forms illustrated in this article is a facsimile, because of space limitations. The figures do, however, accurately represent the question format.*

Figure 1. Example of Teacher Feedback Form



The rating of student interest has undergone several changes. It would have been desirable to obtain an interest rating on each student in order to relate it to their performance and other characteristics. However, such a rating for each activity would be difficult to do with accuracy. For this reason, a rating on the class as a whole was requested.

Figure 2 shows the first two versions of this rating. They can be compared to the final format, shown in item 5 in Figure 1. The first format was difficult for most teachers because it obscured what they felt were significant differences in response. Their ratings tended to be high across most activities, raising some question as to the validity of the information. The second format tried to provide for differential responses. It also offered space to note students representing extremes of response.

In processing the group rating, the five categories were reduced to three: high interest, moderate interest or indifference, and resistance or dislike. It was felt that this would compensate for level of rating and make data from different teachers more comparable. This collapsed scale was used directly on the final format of the feedback form.

The item on equipment problems (item 6) was changed significantly; an earlier format is shown in Figure 3. The earlier format failed to elicit from teachers an explanation of what the specific problems were with equipment; therefore the emphasis was placed on problems encountered rather than on rating the performance of the equipment.

A second section of the teacher feedback form contained questions specific to the particular activity, such as items 12-14 in Figure 1. This section included questions regarding any media utilized in the activity, such as worksheets, slides, films or filmstrips, tapes, games, etc. An earlier form of the feedback sheet attempted to obtain ratings on each type of

(Continued)

Figure 2. Earlier Format for Obtaining Teacher Ratings of Student Interest

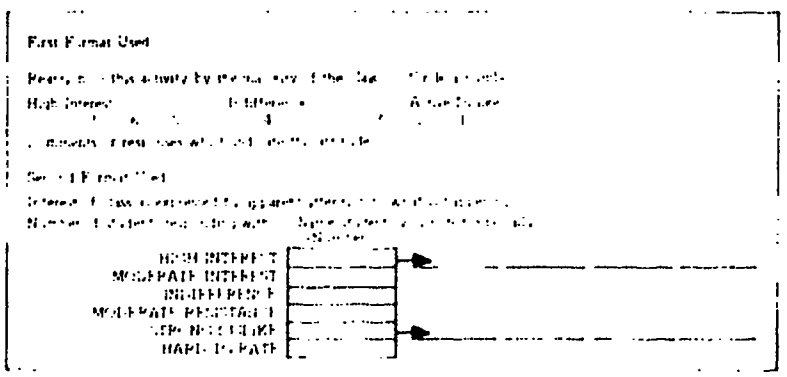


Figure 3. Earlier Format for Assessing Equipment Problems

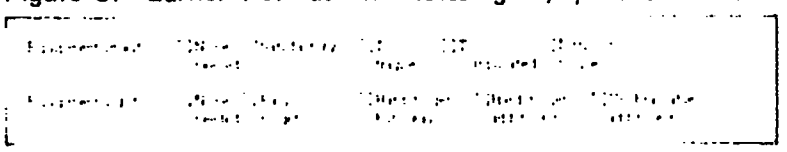
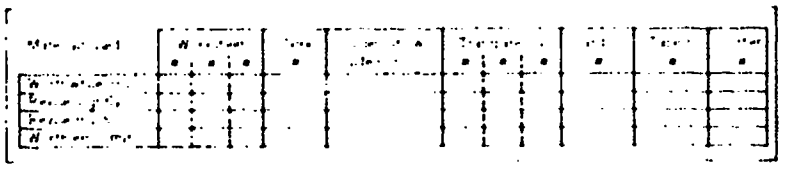


Figure 4. Earlier Format for Identifying Media Problems



media, with comments requested in another place on the form. (See Figure 4.) This format was deleted and specific questions asked because teachers tended to rate some materials and forget others. Also, they neglected to explain the problems and suggest solutions.

A third section of the feedback form was devoted to four open-ended questions. (See Figure 5.) The first item essentially asked the teacher to describe what happened when the activity was used. It has provided an extremely useful dimension of feedback for writers. The earlier format for this item, shown in Figure 6, provided a more detailed set of directions. The check'ist of things to consider was felt to be superfluous by the third year of field-testing and was deleted to allow more room for teacher comments.

Core Reaction Sheet: Following revision of the program into clusters of related exercises (CORES), a second teacher questionnaire–the core reaction sheet--was developed. Teachers were asked to complete this questionnaire, shown in Figure 7, after each of the cores in a Unit. In addition to providing more general reactions to the materials, the back of this form provided a means of updating the class roster and obtaining necessary background information on new students.

The collection of data on student age, sex, ethnic background, and IQ was not the simple, once-a-year task it would appear. First of all, 43% of the IQ scores were far out of date, and retesting was necessary. Often it took considerable effort for the teacher to initiate this testing. Maintaining a current roster of the class was a year-long bookkeeping task for project staff, too, for over one-third of the students transferred out of the class, and others were moved in to take their places.

At various points in the year teachers were asked for additional information, including background data and descriptions of students, ratings of students' abilities, and judgments recorded in the *Student Record of Progress.* Teachers also were asked to administer tests and questionnaires to students and to conduct a number of assessments of students' skills and understandings. See the table on page 32, "Data Collected from Field-Test Classes," for an index to the reports that contain information on these areas.

The mailed ratings of students' abilities required only an hour or so of teachers' time each year The time required for processing this information was also minimal. The *Student Record of Progress* and its associated Clues to Success activities and teacher ratings of students are considered to be an integral part of the curriculum. Therefore, administration and interpretation time spent by the teacher should be considered as part of the overall instructional demands of the materials. Tallysheets were provided for teachers either to summarize data on student performance or to duplicate the information placed in the Record. Each teacher spent approximately twenty-four hours, or three man-days per year, on these. Staff time for preliminary processing and interpretation of this data involved an estimated twenty man-days a year. Of course, developing the assessments, the Tallysheets, and the *Student Record of Progress,* in addition to developing the detailed analysis of the data, required a large investment of staff time. A rough estimate would be 120 man-days for each year of field-testing and for each field test. Part of this time represents developmental costs for the curriculum. A large part is also related to obtaining student data rather than data on the implementation of activities. These costs are discussed later in this article.

Teacher-administered tests and taped descriptions of students required about four hours per teacher per year to supply. Again the costs in terms of development and processing were considerable.

Figure 5. Example of Teacher Feedback Form for Each Activity, Side B

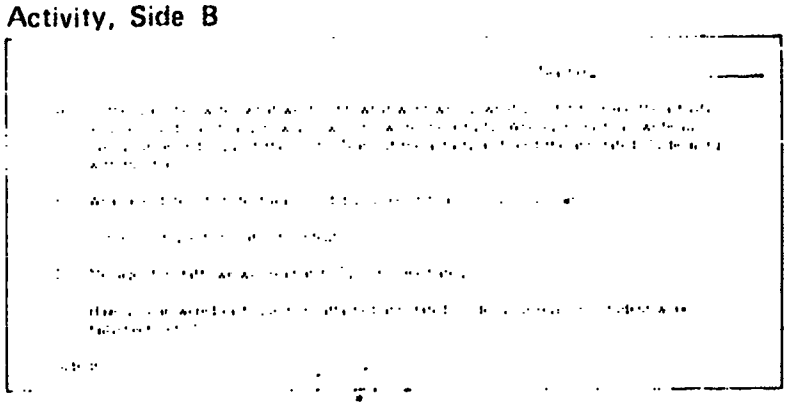


Figure 6. Earlier Format for Side B of Teacher Feedback Form

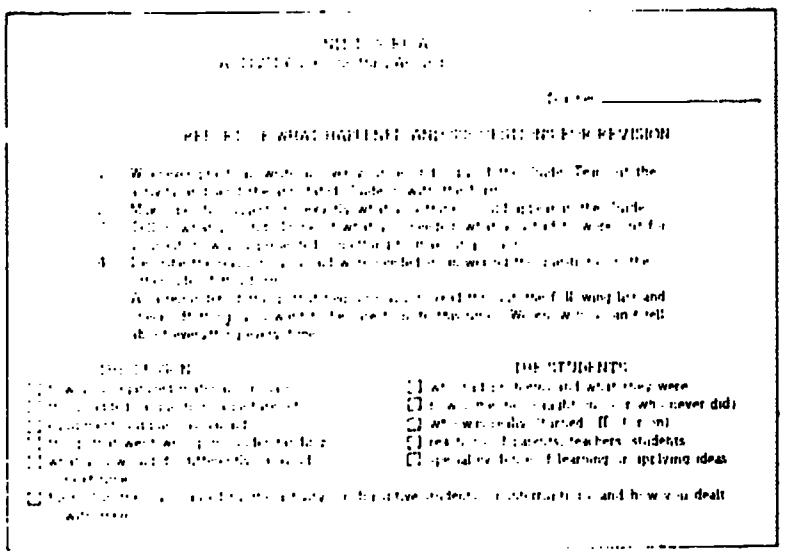
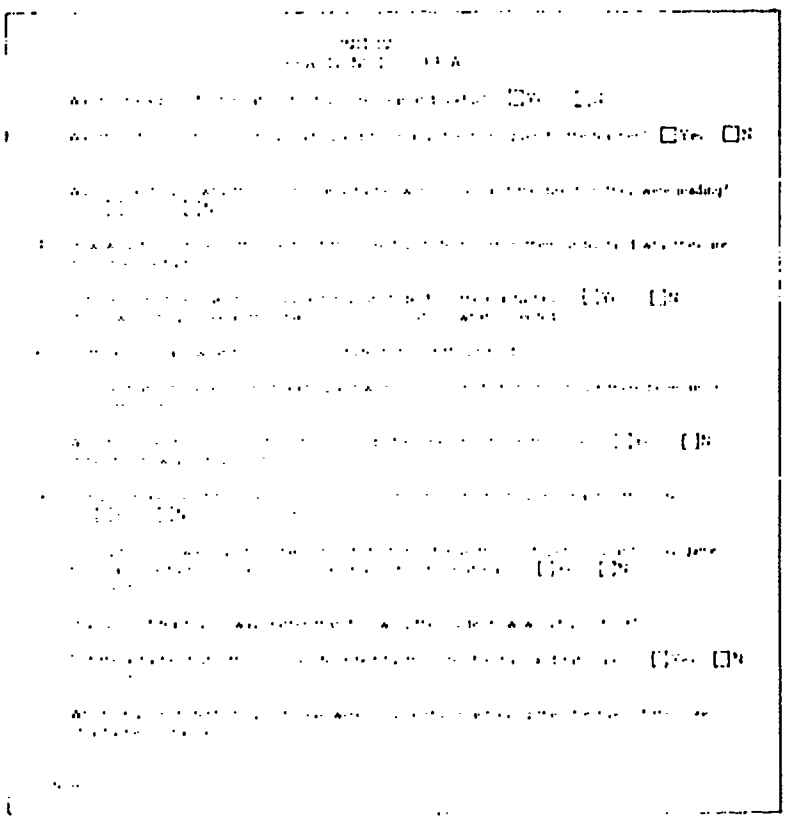


Figure 7. Example of Core Reaction Sheet



COSTS OF DATA FROM TEACHERS

Of all the kinds of data collected from field-test classes, that reported by teachers on activity feedback forms and core reaction sheets was by far the most helpful to writers. To be sure, some of the data used in other studies, such as student functional abilities and performance data, suggested areas of emphasis and entry levels. The teachers' comments, however, told *what* to change and sometimes *how* to change it. What were the costs of delivering this information? Considering only the fourteen teachers involved in the three-year field test of the first experimental version of the curriculum, the following estimated costs were generated:

1. A general format for teacher questionnaire had to be developed, reviewed, revised. The staff time this required was not documented, but probably exceeds twenty man-days for the total development of the two basic instruments.
2. Sections of the activity feedback form specific to each activity were developed for 132 activities. This required review of each activity in an attempt to anticipate problematic and critical elements and to ask about them. A conservative estimate of the staff time for this task is forty-four man-days.
3. Both feedback and core reaction questionnaires were bound directly into the teacher's manual at points of use, and a separate pad of forms was provided so that teachers could retain a working copy. Depending on the amount of description required, each form took from fifteen to thirty minutes to complete. For the complete field test of the first version of experimental materials alone, a total of 1,256 forms were received for processing. An estimate of the time invested by teachers in completing these forms would be about sixty man-days for the total three-year field test. This represents fifty-seven hours of effort for the individual teacher who completed all of the questionnaires.
4. Costs of building a format, printing, and postage were modest. The estimated cost for these two pieces of teacher feedback for the first field-test effort is $500.00.
5. As completed forms were received, they were checked in, screened for completeness (and returned with questions if incomplete), and skimmed for problems of interpretation or difficulties in implementation that required immediate attention. Forms were then filed by activity. An estimated average of one hour per activity was required for these tasks, or nineteen man-days over the total period.
6. A Tallyex computer program was used to summarize teacher ratings across teachers and across activities by teacher. The cost of computer use (key punching, printouts) was about $260.00 over the two years this system was used. An estimated nine man-days were required for coding and processing this information.
7. As feedback was received for each activity, it was reviewed and summarized by a project staff person, and a written set of recommendations was routed to the entire staff for review and discussion. This represents the most crucial step in the process. Estimating conservatively that an average of one-half day per activity was invested in this effort, sixty-six man-days were required. This estimate does not include periodic full-staff reviews of segments of the program or the planning time required to generate specific plans for unit revisions.
8. In addition to reading staff reviews and summaries, writers were requested to read the original teacher feedback as they worked with each activity. A conservative estimate of the time invested in this review is thirty-three man-days for the total first version of experimental materials.
9. A summary of the estimated total cost in time and money for obtaining and utilizing teacher feedback from field trials is shown below. This cost does not include efforts to assess the validity of the data, nor time invested in comparisons and other uses of the data.

	Estimated Man-days	Est. Average Time Value	Estimated Dollar Cost
Teachers' time	60	$60/day	$3600.
Secretarial time	60	$24/day	$1440.
Project staff time	134	$70/day	$9380.
Revision writers' time	33	$70/day	$2310.
Cost of materials and postage			$ 500.
Cost of computer processing			$ 260.
Total time invested: 287 man-days for 3-year test in 14 classes			
Total dollar cost: $17,490.			

When this estimate is extended to the field test of the second version of the experimental materials, the following points must be kept in mind. The revised materials became roughly one-third again as long as the original version. They were field-tested by twenty-one teachers rather than fourteen. Only three of the five units of instruction received a second test, however; Unit II was used in unrevised form by the second field-test classes, and Unit V has not yet been revised or field-tested a second time.

At this point it is possible to estimate the total costs in terms of demands upon the teachers' time for all evaluation requirements. It is estimated that the average teacher spent 91 hours on these tasks, of which 57 hours were related to feedback and core reaction sheets. When one considers that about one hour per activity was required (some spent twice as much time) simply to prepare for teaching the experimental curriculum, the total demand on a field-test teacher was about 140 hours per field-test year. This is the equivalent of an additional work load of one-half day per week, a heavy burden indeed! One can only acknowledge again the enormous contribution the teachers made in the development of materials. Their dedication is indicated by the high degree of cooperation and the high rate of return of needed evaluation data.

(Continued)

DATA FROM STUDENTS

During the first year of field-testing, an achievement test format and pre-post design were used to gather information on student background and entering levels of ability, as well as general information on the students' comprehension of major concepts in the materials. This approach provided a great deal of information used in guiding the developmental effort. This testing format, however, was of minimal use to writers for the revision of materials. Most items developed for such a format tend to assess understanding of major concepts that depend on large sequences of instruction. This was helpful to the staff in determining general and broad activity arrangement, but relatively few items could be tied to individual activities.

The items were also costly to develop and time-consuming to analyze; results were incomplete at the time revision of the first year of materials was conducted. Furthermore, the results of testing were of no utility to teachers in planning and checking their sequence of instruction. Therefore, having obtained requisite baseline information on students and their general level of functional abilities, a different testing format was planned for obtaining performance data in the remaining two years of the field test.[2]

The curriculum was revised to incorporate a number of situational tasks and mini-tests to provide immediate evidence of student understanding or entering ability for the particular subject of instruction. This information was designed to have direct utility to teachers in planning subsequent instruction or review. The results of these short evaluation sequences were also relayed to the project staff for use in analyzing the effects of materials and in planning revision efforts.

This change in methods of collecting field-test data revealed whether students understood concepts and acquired skills at the time they were taught. The data did not provide evidence of long-term retention of information, nor did it allow the calculation of gain scores for individual students, but the technology for adequately doing so is very weak.[3] This was felt to be no great loss.

Instructional assessments were developed for short sequences of activities throughout the entire set of materials. These assessments were treated as program activities themselves, and in many cases they involved practical applications of ideas or actual performance tasks that the teacher rated. Paper and pencil multiple-choice items, such as those that composed the original pre-post tests, represented only one of several formats utilized. In many cases, however, items that were used in the first year of field testing were incorporated into the materials.

For the five units of instruction in the three-year sequence, about 250 assessment items were developed, of which four-fifths were scored for individual students; the remainder were used for judging the class as a group. A tallysheet was incorporated with the activities to use in compiling information on each student and for making ratings of performance. These tallysheets, and in some cases the student worksheets as well, were sent to BSCS and were used as the source of data on student performance.

Several trade-offs were required in electing this method of data collection; among these was a loss of control over the administration of the assessments. Before, the teachers did not know ahead of time what the test items were. Given assessment items as components of the materials, some teachers tended to teach the answers to these items rather than using them to assess learning from other activities. Another trade-off involved the increase of analysis problems due to missing data; by collecting data at multiple points in the year, the percentage of students for whom there was missing data doubled in some classes.

[2] *For a detailed analysis of the first year's assessment of student performance, see interim Formative Evaluation Report 2.*

[3] *Robert E. Stake. 1973. Measuring What Learners Learn. In E. R. House (editor), School Evaluation: the Politics and Process. Berkeley: McCutchan Publishing Corp.*

COSTS OF DATA FROM STUDENTS

The costs of obtaining student performance data in this manner are roughly estimated for one year of field-testing in the first field-test group of fourteen classes. As indicated in a preceding section, the collection and recording of student performance data required about three man-days per teacher per year. (This does not include the class time devoted to this purpose, as it is considered an integral part of instruction.) Thus about 42 man-days of teacher time were involved for the whole field-test group. Another 30 man-days of secretarial time were involved in the receipt and processing of this information. Approximately 120 man-days of project staff time were required in the total analysis and interpretation of this data. (As in all other estimates of costs, the time involved in preparing this series of reports has not been included in the estimates.) The costs of art work, building the format, and production are subsumed in developmental costs. Thus the costs of this aspect of data collection can be summarized as follows:

	Estimated Man-days	Est. Average Time Value	Estimated Dollar Cost
Teachers' time	42	$60/day	$2520.
Secretarial time	30	$24/day	$ 720.
Project staff time	120	$70/day	$8400.
Cost of computer processing			$ 300.
Total time invested: 192 man-days per year for one field-test group			
Total dollar cost: $11,940			

DATA ON STUDENT ABILITIES AND ATTITUDES

The target population for ME AND MY ENVIRONMENT is a group of students whose characteristics and abilities were unexplored in terms of the demands made upon them by this curriculum. Their age made them developmentally different from the population of EMH students previously examined in the development of the ME NOW life science materials. The response to these materials raised questions, however, about the accuracy of ability and performance predictions provided by the field of special education.

Evidence from the first year of field testing of ME AND MY ENVIRONMENT indicated that the data typically gathered by schools on these students has little utility for instructional decision-making. That is to say, IQ, age, sex, and ethnic background explain little of the variance in performance of these children—at least in these experimental materials. (See interim Formative Evaluation Report 2 for details on these findings.) Therefore, concurrent with the testing of materials, an attempt was made to identify functional characteristics of students that were meaningful to teachers and useful in planning an instructional program. Teacher ratings of selected skills and attitudes of students, described in the previous section and in Report 2, were one aspect of this search. The cognitive development and problem-solving abilities of these children were also studied. The refinement of an instrument to assess these dimensions is described in detail in interim Formative Evaluation Report 4.

We are indebted to William M. Gray for the research underlying the paper and pencil assessment of Piagetian levels of cognitive development. This instrument was revised and administered yearly during the course of field-testing. An extensive validity study has accompanied use of the instrument. This effort has included test-retest studies, follow-up interviews of student samples to obtain explanations for their answers, and readministration of items using the traditional Piagetian format. The data generated by this instrument, along with the teacher ratings, have been used to suggest modifications and additions to the curriculum materials.

The cost of developing the cognitive development/problem-solving instrument has been relatively modest, in part because of the availability of previous research. Approximately ten days were used in developing and revising items. Total time required for administration and scoring was approximately fifteen days per year, plus another fifteen to twenty days for validity studies. An additional cost of several hundred dollars was required for an external analysis of the structure of the instrument.

Four questionnaires have been developed and administered to all students. One dealt with student awareness of and attitudes toward the use of alcohol, tobacco, and drugs (Unit II). The other three questionnaires assessed attitudes toward science activities in Units I, IV, and V. Each questionnaire was administered in less than one class period, and no more than two were responded to by the same students in any one year. These instruments required very little time to develop or score.

Five structured interviews were developed for use with stratified random samples of students. In addition to these, a number of interviews related to validity studies were conducted. Informal interviews by observers and staff were conducted occasionally in conjunction with class observations, without using a sampling plan.

The structured interviews involved a variety of costs and trade-offs. They were conducted by project staff during site visits; they required about thirty minutes per student, and usually four students per class were interviewed. Because time was limited during site visits, interviewing was done in lieu of additional classroom observation. Problems other than cost were factors in collecting this data:

1. Each interview protocol required ten or more man-days to develop.
2. There was seldom sufficient time to field-test and revise the instrument prior to use.
3. Standardization of procedures among staff members and use of appropriate interview techniques were difficult to maintain. Some data were invalid as a result.
4. Tape recordings were made of each interview. Some interviews were lost because of faulty equipment and errors in recording. Some tapes were quite difficult to understand because of background noises and lack of appropriate school settings for recording.
5. Tapes were listened to and coded by the interviewer as well as by another coder. This process required about two times the length of each interview to complete.
6. Typed scripts of some interviews were made. Each script required about two times the length of the tape to type.
7. The interviews had a considerable impact on staff members' judgments of the students and the program. Their impressions were biased by the partial sample they were directly exposed to. It was difficult to obtain a perspective based on the entire sample.
8. The interviews were generally collected over a time span of two to three months, making comparability of results difficult. Analysis and interpretation were also delayed because of this time lag.
9. The sheer length of time required in listening to the total set of tapes created the sense of a huge burden of data to process. The richness of this data also made interpretation and summarization quite difficult.

A single interview, sampling about 25% of the student population, required a total of approximately thirty-five to forty man-days. In light of questions regarding validity of the data and distortions acquired by the impact of partial data, this cost seems excessive.

In addition to this variety of data obtained from students, samples of worksheets, posters, and other products were requested at selected points in the instructional sequence. These were studied to identify problems and determine the level of understanding exhibited by the students. *(Continued)*

DATA FROM OBSERVERS

Site visits to field-test classes by project staff were felt to be a critical source of data. This data was a necessary accompaniment to the extensive feedback from teachers, for it enabled data from each site to be qualified on the basis of a number of context variables. It probably had almost as much impact and influence on the revision of materials as teacher feedback.

Four classes in the first field-test group were selected within a one-hour drive of the project offices. All staff members visited these classes periodically during the year. In addition, during the first year of field testing, each of the ten out-of-state sites was visited twice by a member of the project staff. Each of the two-day visits was primarily spent observing the students in class, both during science instruction and during other periods of the day. Interviews with teachers and students and discussions with administrators were also held. The visits provided teachers with an opportunity to clarify questions, to discuss planning and problems, and to resolve any difficulties encountered in providing feedback. These site visits provided the monitoring necessary for judging teacher fidelity to the instructional manual and to the intent of the materials. The visits also provided considerable insight into the context for instruction. A written report was filed for each visit that described the class, the context, and the judgments formed. In addition, notes regarding the activity being taught and ideas for revision were filed.

In the second year of field-testing, with two sets of field-test classes in operation, the staff could make only one two-day visit to each site during the year because of the twenty-eight out-of-state sites participating in the program. In the third year of field-testing, twenty of the out-of-state sites were visited once. With each out-of-state site visit costing an average of $300.00, the total investment in this form of data collection for the three-year period was over $20,000. This figure does not include the 200 man-days of staff time (including travel time) required for the visits.

In addition to these visits, four classes located near the project offices were observed during every science period throughout the first two field-test years. Four graduate students observed the classes and wrote up detailed reports for every day of observation. Observers with differing background experiences were selected. One observer was a graduate student in sociology who had worked with disadvantaged groups, a second was specializing in urban psychology, a third had experience as a special educator, and the fourth observer's background was biology.

This full-time observation provided feedback from a different perspective than the teacher's, and it allowed a reliability check on the data teachers reported. A variety of procedures were tried in an attempt to discover which kind of information was most useful for revision. Eventually, the observers were asked to turn in a report for each activity that focused primarily on three questions:

1. What evidences can you list of the successful implementation of this activity (or of problems of implementation)? [This item has to do with the mechanics—whether the planned activity worked, if the strategy facilitated doing it, if there were organizational problems, etc.] Why do you feel these are evidences of success?
2. What evidences are there that the teacher understood (or misunderstood) the intent of this activity? [Some appropriate details include the questions asked, the mode of response to students, capitalizing on events, etc.] Why do you choose these details?
3. What evidences did you obtain that students comprehended (or did not comprehend) the concepts in this activity? Why do you feel these are evidences of success?

The observers were expected to document their answers with specific events and details. They were also asked to justify their choice of evidence by explaining the standards and criteria they used. They were expected to read the activity in advance and to decide what information would be needed in answering the questions. Then they planned to do whatever was necessary to gather the data (interview students or teacher, watch for an event, rate student performance, etc.). In addition, the observers noted unexpected or unintended effects of the strategies and made suggestions for revisions.

The costs of observation were $40.00 per day for about three hours from each observer. Approximately sixteen to twenty hours of training time with the total project staff was provided before initiating the observations. At turning points in the instructional sequences, observers met with project staff for review and debriefing. These meetings consumed about twenty-four hours during the course of a year, or twenty-seven man-days of time. Observers' notes were included in staff and writer reviews of each activity. The total cost of this activity per year was approximately $11,500. In spite of the valuable data derived from this source, observations were discontinued for the third year of field-testing because of budget constraints.

In addition to the kinds of data collection described in this article, other studies have been initiated but discontinued for a variety of reasons. These incompleted studies include such things as taped student discussions from test classrooms, videotapes of instructional sequences, and student interviews on several topics. Problems that led to discontinuing these studies included the mechanical difficulties in obtaining interpretable data, technical problems in obtaining an adequate and valid sample, and time and money constraints. These interrupted or aborted studies will be discussed in a later report in this series. ■

organizational procedures and decision criteria

The central dilemma of most evaluation efforts is the desire to ask more questions than it is possible to answer with resources that are available. A twin dilemma is the desire to collect more data related to the questions asked than can be processed and utilized. This combination of overextension and data overload can destroy the effectiveness of the most elegant evaluation design. In formative evaluation especially, the efficient processing and delivery of results is of paramount importance if the evaluation is to make a significant contribution to the development of a program.

Presuming that these central obstacles to the utilization of evaluation findings are overcome, by what magic are the results incorporated into the developmental and revision process? The mere delivery of results does not insure their use. It may often be quite uncertain what implications should be drawn. A plan for delivery of evaluative evidence that includes some direction for its use can increase the ultimate payoff of the evaluation.

The descriptions in this section are not presented as solutions to these dilemmas. They do indicate how these problems were addressed in this evaluation. Because processing and delivery are so critical, it is believed just as important to report the procedures used as it is to document any other stage of the evaluation effort.

PLANNING AND REVIEWING THE TASK

At the end of the first year of this curriculum development project, a detailed estimate was made of the time required to perform the various tasks of development and evaluation. By that time it was possible to make much more accurate estimates of the time various tasks would require than could be made at the beginning. Based on the one year of experience and the original evaluation plan, the following estimates were made for the second year of the project:

1. Site visits to 34 classes	286 staff days
2. Teacher orientations for 2 field tests	65 staff days
3. Preparation of instruments	86 staff days [an underestimate]
4. Analysis of all data	487 staff days
5. Advisory committee meetings and consultations	126 staff days
6. Summer writing conference	200 staff days
7. Materials review, editing, and revision	120 staff days [an underestimate]
8. Staff meetings and office routine	100 staff days
9. Workshops and presentations	180 staff days
10. Additional unexpected tasks and projects	200 staff days
Total staff time estimated	1,850 staff days

These estimates are summaries of the more detailed analyses of each task.

The estimates indicated that 7.7 full-time staff equivalents would be needed to perform the tasks. For 4.5 staff members, the plan was obviously unrealistic. Part of the great difference between plans and resources lay in the modifications that had been made in project plans:

1. The curriculum had been restructured as a three-year sequence rather than the two years of materials originally planned.
2. The decision had been made to run field trials of the revised materials simultaneously with the continued field test of the second year of new materials.
3. An agreement had been made to complete the second revision of Unit I for commercial release by September 1973, including provision of camera-ready copy.

As a result of this review, reductions in data collection plans were initiated. Visits to out-of-state sites were cut from two to one during the year. Pre- and post- achievement testing was discontinued. An additional member was added to the project staff. These changes reduced the burden of tasks and increased the manpower equivalent to within one-half a full-time staff member of the estimated demand. The budget was reviewed to identify funds that could be used to augment the limited amounts originally allocated to evaluation. And finally, efficiency in operation was considered even more essential to ease further the burden imposed by what was felt were necessary steps for adequate evaluation.

STAFF RESPONSIBILITIES

The administrative structure of the staff emphasized a team approach to all tasks. Various staff members accepted ultimate responsibility for various functions of the project—director, writing supervisor, materials and media, evaluation. Roles and tasks, however, were not reserved strictly for one staff member or another. Instead, each staff member had the

(Continued)

responsibility of writing and critiquing activities, helping design and package materials, and performing evaluative tasks. Decisions were reviewed and agreed upon by the full staff. Much initiative and cooperation was exercised by the various staff members in seeing that all tasks were carried out.

Evaluation was intentionally defined from the beginning as a role of each member of the staff. Indeed, the staff was the major audience to whom the results of evaluation were directed. In order to serve this audience and to incorporate findings into revisions, it was reasoned that the results would be better understood and accepted if the full staff had an integral part in conducting the evaluation. Furthermore, the findings had to be utilized before final processing and formal reports could possibly be written.

The staff members themselves were viewed as evaluative instruments. Their involvement with the data–the site visits, class observations, student interviews, teacher feedback, results of student assessment–all were internalized, discussed, and generalized. Such involvement in the evaluation process served to program the staff with ideas and concerns that then formed a basis for guiding and reacting to the writers' efforts in revision and further development.

In this context, evaluation tasks were seen, along with the other tasks, as everyone's responsibility–not given a low priority, as is often the case, and left until last to do. The other tasks on the project were also viewed as the responsibility of all. This view cast the evaluator into new roles, too.

The evaluator was involved in the developmental tasks of the project and had an equal input in critiquing and contributing to the design of the program. The effect of this shared responsibility was to maximize trust in each other's abilities, to broaden the understanding of the framework and basis for decisions and the direction of effort, and to provide an open exchange of views. It was possible to criticize or modify one another's writing and interpretations without rancor or hurt feelings. The ego involvement and defensiveness that accompanies exclusive territorial divisions of roles was absent.

KEEPING TRACK OF DATA

Coordinating the efforts of four to five people, keeping informed of the progress of the field tests, and being able to find data are basic conditions for efficient work. Some of the simple procedures used to accomplish these ends include the following:

1. A bulletin board chart was posted daily for all feedback received from teachers, indicating the date and the activity completed.
2. Staff members shared the task of scanning teacher feedback and responding to teachers. Feedback of special interest was routed to all staff.
3. A schedule for classroom observation by staff and observers was kept up to date.
4. All contacts with teachers or classrooms were briefly described in writing and routed to the staff.
5. A copy of anything pertaining to a specific activity was filed in a folder for that activity.
6. Anything pertaining to a teacher or class other than activity information was filed in a folder for that teacher.
7. Newsletters were sent periodically to teachers to keep them up to date on each other's progress and to inform them of corrections, helpful extensions of activities, and suggestions for helpful feedback.
8. Personal letters to teachers were written whenever any questions arose about feedback or problems about instruction. Incomplete forms were returned to teachers for completion.

PROCESSING DATA

A set of procedures was agreed upon for the shared responsibility of summarizing and interpreting the various forms of data being collected. The following guidelines were used for processing several different kinds of information:

Data on Use of Activities: Each staff member was assigned specific activities to review. All data on file by a predetermined cutoff date were included. This assignment was made early enough for the staff member to review the activities and to anticipate any problems or additional information necessary in time to request it of teachers.

It was not realistic for staff members to observe each activity they summarized. However, they were requested to observe at least one class session of any activity that could present problems or special opportunities to judge student learning.

The task of reviewing an activity included a number of steps:

1. Teacher responses to the first section of the feedback form (items 1-11) were coded and computer-processed, using the Tallyex Program. These results were scanned to note any areas rated unsatisfactory.
2. A review was made of teacher comments to each item on the feedback form, with a summary of ideas held in common and/or unique suggestions judged worth incorporating.
3. Teacher comments were qualified by ratings of fidelity to the teacher's manual. Observer and teacher feedback were compared. Responses were also sorted by the context of the class: age, ability level, location.
4. A preliminary review of the student assessment items was made, including concerns about the validity and functioning of the items.
5. Student worksheets and/or other products were examined and areas of concern noted.
6. A written summary of the activity was prepared, including the following sections:
   a. General summary and teacher recommendations
   b. Analysis of specific questions on the feedback form
   c. Media concerns
   d. Anecdotes
   e. Reviewer's recommendations

This summary varied by the activity and by the person writing the review. See page 29 for an example of a review. When all activities for a core had been summarized, the core was reviewed. Steps in this review involved:

1. Tables were prepared, summarizing ratings across all activities in the core for each of the Tallyex ratings (first 11 items) on the feedback form. This provided a correction for the level of ratings, identifying activities rated lowest in the group.
2. The Tallyex data were rerun by teacher to look for individual variations in judging the activities in the core.
3. The student assessment items were computer processed, using an item analysis program. Conclusions were drawn for groups of items as well as individual items. Results were looked at by class as well as by total group.
4. Teacher comments were collated for each item on the Core Reaction Sheet.
5. The general pattern of results across this sequence of activities was examined with an eye to expansion or restructuring of the sequence.
6. A staff meeting was held after each person had reviewed these summaries; reactions and recommendations were aired. Occasionally observers would participate in these sessions.

Data from Student Interviews: Each staff member was involved in the final stages of interview protocol development for each instrument used. Each component of an interview was referenced to the purpose for including it and the plan for analysis of it. The scale to be used for coding responses was also generated at this time. A scoring form was developed to allow interviewers to code responses during or immediately after each interview. This degree of development of the instrument enabled interviewers to probe responses sufficiently to obtain the degree of clarification necessary for scoring.

For most interviews, students in all test classes were assigned random numbers and four children plus alternates were selected from each class prior to visiting the site. In the second and third years of field-testing, the sample from each class was stratified into those students who were continuing from a previous year of involvement in the program and those students who were new to the program.

After each classroom set of interviews was collected, another staff member would audition each tape and code responses. This score was compared with the interviewer's score and differences resolved through discussion. No typescripts of interviews were prepared unless the student comments provided additional insights into the program or the scoring protocol required collating actual responses of various kinds.

No attempt was made to use the results to make judgments about individual classrooms. The nature of the sample did allow conclusions to be drawn about the field-test group as a whole.

Data on Student Performance: In each year of field-testing, a log was prepared showing each item of data being obtained on each unit of instruction. These data included information for which a score could be assigned for each student, information that was not appropriate to score for individuals or which was not collected for all students, and specific questions posed for each activity. This log indicated which student worksheets would be retrieved for full classroom groups, and when any other student products were being collected. Such a log, combined with the daily check-in of material and a supporting statement written for each instrument in use, provided a fairly complete "map" of the evaluation effort for the staff.

The log of data collected from students was reviewed by the staff prior to analysis of results from the second field-test year. For each item the staff made a judgment of how central or peripheral it was to the purposes of the curriculum. On this basis it was possible to weight some items more heavily than others. A second judgment involved what would constitute an acceptable level of success on each item. This was not a prediction of how we thought students would perform on the item. Instead, it took into account how many times the concept or skill had been presented to students, the cognitive level of the item, and the centrality of the item to program purposes. On many items, partial credit was allowed for some responses. Some items had more than one acceptable answer, and extra credit was allowed for recognizing more than one right answer. Some standards were as low as 25%, others as high as 90%. Standards were also set for items grouped as subtests. Recognizing the arbitrary nature of these standards, the staff felt that they did provide a reference point not prejudiced by the actual results, and also that they did prepare us for interpreting the results.

As results came in, they were examined class by class for the indicators of credibility noted in the previous article on student data. Most passed this screening, were coded for machine keypunching, and analyzed, using an item analysis computer program. Results were then compared with standards, and dimensions falling below standard were identified as revision concerns.

Some forms of data are more troublesome to process than others. Typically, narrative information can become overwhelming if allowed to accumulate. It can be even more difficult to deal with if it is desired to categorize comments and quantify them in some way. Since the value of the remarks was in their content rather than in their variety or in the proportion of comments of various kinds, no more sophisticated reduction was needed than the reviewer summaries.

Data recorded on tape proved to be the most difficult to process for this project. The first problem was that it was accessible only by listening from start to finish, or by having a typescript made. Second, it generally arrived in large bunches. The psychological effect was that of a never-ending burden, that the listening and the coding of information would go on forever. This led to avoidance behavior and to reducing the priority for processing this data, with subsequent delays in its full utilization.

CRITERIA FOR DECISIONS

Ultimately, the decisions facing curriculum developers reduce themselves to these few:

1. To retain the materials in their existing form
2. To retain the materials with minor revisions
3. To retain the materials with major revisions
4. To expand or extend the materials
5. To add new materials
6. To revise program objectives and add, modify, or drop materials accordingly
7. To drop the materials

Some of these decisions are made on a logical basis. Some are made intuitively. Some are based on a variety of information, part of it derived from evaluation efforts. To increase the proportion of decisions that are based on data from evaluation, it is important to specify what standards apply to each component under study.

These criteria are unique to specific situations. They are important primarily as a predetermined plan for interpreting the evidence. Criteria are reflected in the plans for analysis of each instrument used in the study. The original choice of questions was based on specified criteria. In that case, brainstorming had generated over 100 questions relevant to the study. Using the Delphi technique, the experience of a diverse group of judges was used to reduce the number to a few that were viewed by all as central. These questions then became the criteria for choosing kinds and sources of information.

As implied in the preceding description of standard-setting for student performance items, many of these standards may be arbitrary. They reflect the opinion of informed judges and are based on the demands of a situation uninfluenced by actual results. Other standards, such as those used to weigh teacher feedback, rest upon a series of judgments about each teacher's credibility and fidelity to the intent of the materials.

The way student performance is combined with teacher feedback and a variety of other information in revision decisions remains intuitive. Some developers and writers place more weight on outcomes; others attend more to processes and transactions. Whatever the predominant cues, the revisor is asked to justify his or her decision in the light of all the accumulated information. ■

the case of the sugar classroom versus the medicine classroom

Nothing written in these evaluation reports really captures the climate, mood, and feel of the classroom. Even a classroom visit cannot reveal the full impact on students of life in the school and in science classes; only a series of visits could show that impact.

The meaning of school still cannot be understood clearly without becoming acquainted with the students in their rooms and spaces at school, and in their neighborhoods and homes. Only with this kind of knowledge is it possible to view the curriculum under test in its proper perspective, and to begin speculating on its meaning and outcomes.

Such detailed views of the classroom and the lives of students were not included in the evaluation effort, although observations and informal interviews recorded some of them. Many of the personal details gained through visits and discussions with full-time observers had an impact on the staff's perspective of the classes. Thus some of the classroom experiences did become a part of the judgmental base for interpreting feedback.

The feedback—data that was collected as a central focus for study—was chosen to represent the concerns of the developer, not to portray the reality of the classroom. While such feedback reveals the effects of the materials and reflects some elements of student performance and response, it is possible that it may be peripheral to the full impact of the program on students. The curriculum tested did not remain distinct and separate, but meshed with the various "programs" pursued by the teacher and with the characteristics of students and school to form a different amalgam in each setting.

It is easy for developers to shake their heads and object that the teacher does not understand the intent of the materials—or to nod and point out how well the curriculum "works when it is used right." To say such things is to impute far too much to the materials. Many of the transactions in the classroom are far beyond the power of sheer materials to affect. Indeed, they are beyond the skills of most developers to change through direct intervention. Many dynamics of the classroom might serve to annotate both good and bad results of the kind typically reported for "field tests." What follows are a few of the features of the classroom that should qualify the data of this study.

Two classrooms were chosen as contrasts. The first year of field testing is the time interval involved. These classrooms are real, not composites; however, the selection of incidents to draw a contrast between classrooms reduces the accuracy of what is intended to be a representative case study. The intent of the comparison is to dramatize dimensions of instruction that should concern developers: for this purpose, the distortion, although unfair to the teachers, is believed justified.

THE CLASSROOM SETTINGS

There were many points of similarity as well as difference between Class S ("Sugar") and Class M ("Medicine"). The personalities of the teachers (hereafter called "Sugar" and "Medicine") were pleasant, but quite different. They were both in their second year of teaching and were working with seventh-grade EMH children in semidepartmentalized situations. Both cared about their work and about their students. Both were interested in outdoor sports, such as hiking and skiing, and tried to arrange for their students to enjoy these kinds of experiences.

The schools and students differed in many ways. Both schools were targets for Title I programs, but one (Medicine) was located in a low-income, inner-city area of rundown homes, apartment buildings, and welfare housing. Residents were predominantly black and Chicano. The other (Sugar) was near a commercial section of a lower middle-class suburb, with residents predominantly white blue-collar workers. Both were in large metropolitan areas.

THE STUDENTS AS A GROUP

The students may be viewed from a variety of perspectives. The Sugar classroom had sixteen students compared to fifteen in the Medicine classroom. Both classes were two-thirds boys. The mean age in both classes was 13.7, with most of the students 13 years of age. The mean IQ level was seventy-five and seventy-six in the two classrooms, and both had six students with IQ scores of eighty or above. The ethnic composition of the two classrooms was quite different. The Sugar classroom had twelve white, one oriental, and three Chicano students. The Medicine classroom had one white, ten black, and four Chicano students.

Measures of performance based on instruction in science often do not reveal basic differences, so covariance analysis was used to adjust for differences in entering abilities of students. Covariates in this adjustment included IQ, problem-solving ability, and ability to follow directions. Significant differences in performance were found among the ten field-test classes; however, the Sugar and Medicine classes had the second and third highest mean scores among the ten.

Over the course of a year one student was dropped and one was added to the Medicine classroom; two students were dropped and two were added to the Sugar classroom. From year to year, however, the attrition rate is higher at the Medicine school. By the second year of field-testing, half the class was new. Five students moved away, one was moved out of special education into the regular program, and one became pregnant and dropped out of school. In the Sugar school, one student moved away and one was moved into the regular school program.

Attendance at the two schools provides another contrast. Based on absences during instruction in Unit I, six students were absent more than 10% of the time in the Medicine classroom, with three of these out half the time or more. Three students in the Sugar classroom were absent slightly more than 10% of the time.

All of these differences contribute to the contrasts noted below, but they don't explain them away. The reader is urged to consider the *implications* of the conditions and transactions that are described. What do they imply for children; what do they say to developers?

THE STUDENTS AS INDIVIDUALS

Space does not allow the description of each of the thirty-one students in these two classes. Many missing details would also limit such a description. Several students from each class are described in interviews, teacher comments, and records provided by the school. Names have been changed for obvious reasons.

(Sugar Classroom) Jerry is a tall, good-looking white boy. His home environment is good. He lives about two miles from school and rides the bus. The family (mother and father, two sisters) has lived in the area about four years. Jerry was wearing a cast early in the year because his locker was slammed on his fingers during a fight. Jerry says he likes home better than school because he can watch TV and go outside with his friends. At school he likes science, math, and art.

Jerry turned fourteen during the school year. His measured IQ is eighty. Late in the year it was discovered that he has a total hearing loss in one ear and a partial loss in the other, a condition that has existed for some time. This explains the teacher's description early in the year: "Jerry is probably in the upper part of the class. Sometimes he doesn't pay too much attention, but when he does he has the ability to follow directions and to complete his work. He is well liked by the other students and seems to have no physical or emotional problems. He is also very polite and gets along well with adults. He sometimes has to be reminded to hand in his assignments or to do his work. He tends, quite often, to doodle or draw pictures rather than to do his assignments."

At the end of the year the teacher said: "Jerry's behavior was extremely erratic. He was a perfect boy at times; at other times he would sit and draw pictures instead of listening, and be very belligerent. Now the hearing problem explains a lot of this. He would not ask you to repeat the question; he would just mumble something and turn his head away. All three of his teachers thought this was because he did not like school and did not want to perform. But now we feel it was probably because of his hearing problem. Jerry does well in school. He is an extremely social boy. He likes activities, mechanical things that require him to use test tubes and beakers and so forth. He did quite well in science." (See anecdotes from science class for another view of Jerry.)

(Sugar Classroom) Cami is a quiet, tall, good-looking Chicano girl, very personable, yet shy. She rides the bus several miles to school. At home she lives with both parents, three sisters, and two brothers. She says she prefers school to home and doesn't like anything at home, but she does not give a reason. At school, science is her favorite class. Cami turned fourteen during the school year. Her measured IQ is seventy-seven.

Early in the year her teacher commented: "Cami does her work well. She follows directions. She knows the answers, but because of her shyness she does not speak out much in class. She gets along well with others and is not afraid to defend herself if teased or accused of something. She functions in the middle of the group—about second- to fourth-grade level. She seems to be well adjusted emotionally and her home situation seems to be quite good. I believe she sews quite a bit and has quite a few responsibilities at home."

At the end of the year, the teacher's view had changed: "Cami was absent a great deal in the second half of the year. Cami's parents have not been to a conference. I tried to get in touch with them to make a home visit, but could not do that either. When Cami was here, she had little difficulty understanding what we were talking about; she followed directions well; she had a lot of social awareness; and she carried out the science activities well. Of the girls, Cami was one of those more interested in science, but she didn't show a great interest. She is a leader among the girls and tends to be a little bit catty or bossy at times."

(Sugar Classroom) Sandy is a stocky, red-headed lad who always appears cheerful and interested in participating. Sandy's family (both parents, two brothers) recently moved into the area and live several miles from school. He says, "I like school a little bit, but home the best. We got two birds at home. One belongs to me and one is my brother's. We got two dogs, too." At school he likes art and science. His home environment is normal and his parents are very cooperative. Sandy turned fifteen at the end of the school year. His measured IQ is seventy. He has been in a special education classroom for four years. Early in the year his teacher gave this description: "Sandy functions quite low. He works best when there is somebody next to him, encouraging him to do assignments, explaining to him what he is to do, forcing him and helping him follow directions. He will not participate in anything that has to do with reading, spelling, or verbal tasks.

(Continued)

He is almost a nonreader. He has not had n success experiences at all. He is dependable and very, very eager to help. When the children in the class are disruptive, he encourages them to be quiet, though he is a little bit rambunctious himself."

At the end of the year, the teacher commented "Sandy has come a very long way from the beginning of the year. Now he is able to read at perhaps a first-grade level. He can write with help and has come to the point where he does not just sit and wait for somebody to help him. In a state-wide art contest held by the governor's committee on the handicapped, Sandy entered a papier-mâché animal he made and won honorable mention on it. He even had an offer from someone to buy it. This raised his ego quite a bit, not only in art class but in a lot of other classes. He seemed to enjoy science very, very much and had a good ability to reason things through logically rather than just shouting out an answer. He did well in science; he liked it, and kept asking when we were going to have it."

(Medicine Classroom) Jose is a very pleasant fourteen-year-old Chicano boy. His home situation seems normal. Spanish is spoken in the home. He is slight of build and has an innocent heart murmur. His measured IQ is seventy-eight.

At the first of the year, his teacher said, "Jose frequently visits with his neighbors, so he is sitting in the front of the room. That cuts down some on his talking. He can follow directions, but I have to repeat them four or five times before he gets them. He is not sure of himself; he has little self-confidence. He participates in class—sometimes he has good answers, sometimes not. He doesn't think out his answers, just blurts out the first thing that comes in his head. He has good social adjustment and can work in a group. His current academic level of functioning is probably second or third grade. He has an average amount of curiosity, an average amount of enthusiasm, and an average attitude toward school. He doesn't usually accept the responsibility of doing outside tasks."

At the end of the school year, the teacher commented: "Jose transferred a month ago, as his family moved out of the district. He was a good student, a friendly child. He didn't get along with the class too well. He was an individual—not really a loner, but he didn't participate in group activities. He did get involved; he did enjoy science. He was excited about many of the activities. I was sorry to see Jose leave." (See anecdotes from science class for another view of Jose.)

(Medicine Classroom) Fern is a big, black girl—not fat, just big all over. Her parents and two sisters and a brother have lived seven blocks from school for six years. They seem to be a close, warm family. Fern is quite proud of her older brother and talks about him frequently. Fern says she likes school best because she doesn't like to do the housework at home. At school, she says, "I like science, math, home ec, and I like to play around in the halls." Fern turned fifteen during the year. Her measured IQ is eighty-two.

At the first of the year, her teacher gave the following description: "Fern is probably the highest functioning student I have. She is very verbal, wants to give the answer, thinks out her answers, and usually has the right answer. She has not been too prompt in getting her homework in. She usually turns it in two or three days late after she realizes she will be punished if she doesn't. She likes school and is one of the leaders in the class. She usually keeps the other kids quiet when something new is going on that she wants to learn about."

At the end of the year, the teacher said: "Fern talks real tough, but is a real sweet person. She is popular in the class because she can get along with everybody, yet because of her size she can maintain quite a bit of authority. Whenever I have to leave the class for a few minutes, I put her in charge. Fern responds to science quite well, and she does the activities. She is a good student: she isn't a discipline problem. Fern participated in Special Olympics and got first place in the softball throw. During the summer, she became pregnant and dropped out of school."

(Medicine Classroom) Marvin is a tall, good-looking black boy who has lived five blocks from the school for five years. Both parents, a brother, and a sister live at home. An older brother is in college. He says he likes home better than school, and he likes to play basketball. This attitude is supported by the fact that he was truant 47% of the year. His teacher says that his parents let him stay home and that she often sees him riding his bicycle if she passes his house during the day. At school, Marvin likes reading, gym, and arithmetic. He is fourteen and has a measured IQ of eighty, although the teacher feels this assessment is high.

At the first of the year, she said: "Marvin has to have directions repeated over and over and over. He can't pay attention during class instruction, but after watching me do the work four or five times, he finally understands what he is supposed to do. He also has to be reassured. After every single question he will come up and ask, "Did I do this right?" He likes to bully, shove and play around, swipe kids' pencils, and things like that. Whenever he is criticized he gets very defensive. You can't get through to him; it's almost as if he blocks it out. He never feels he did anything wrong. He is curious and will get enthusiastic about what he wants to study."

At the end of the year, the teacher repeated many of the same observations: "Marvin is not sure of himself. When he does English worksheets, he'll do one question, come up and see if he is doing it right; do the next question, come up and check it out, and so on. He is absent quite frequently. When he is in school he is generally pretty well behaved. He does little petty tapping, kicking, talking, and hyperactive things that most of the class members do. His involvement with science is quite good when he is here. He likes many of the activities."

STUDENT RESPONSE TO MATERIALS

The "Sugar" teacher feels the experimental science materials are well suited for her students. She thinks most children enjoy the activity part of science more than sitting down and discussing it. At the beginning of the year the students had a great deal of difficulty in following directions and carrying out activities because they had never been given this freedom before. At the end of the year the teacher felt she could distribute materials, give students directions, and trust that there would be no problems with potentially dangerous materials.

The teacher did see a gap in the materials for her slowest children; she felt the pace was too quick for them. But if she slowed down, the majority of the class would have been bored. There were many things that went completely over several students' heads.

The "Medicine" teacher feels students are really turned off to most educative sources, although they do like to read sometimes in their reading books. She thinks they don't like the science experiments much better than anything else they have, but she thinks they would *hate* science books that might be used in place of the experimental program. They did like many of the activities, she feels, and she explains that there are

(Continued)

many things her students outwardly appear apathetic towards that they are really enjoying. They usually hate to get started on anything, but once they get going they enjoy what they are doing. However, when they are doing activity-oriented things, this teacher feels she cannot let them have total freedom to explore because they lose sight of what they are doing or start messing around. She says, "I have to discipline them and that may stifle their interest in the activities somewhat. The students are easily bored and if anything starts to drag, they get hyperactive."

GENERALIZATIONS ABOUT THE SUGAR CLASSROOM

This teacher approached most lessons enthusiastically, attempting to dramatize the activity and catch the students' interest. She responded to the rhythm of the class, recognizing when to change pace, have a break, or move on to a new activity. She seemed to have internalized the program and was conscious of the intent of each lesson, capitalizing on unplanned opportunities to develop further understanding of concepts.

Though she set clear standards for behavior, she was tolerant with her students and seemed to recognize they were children who were full of energy and who would occasionally misbehave. She was warm and responsive and would stop to listen to her students. She recognized their moods and feelings and could empathize with them. She gave encouragement and praise often.

The pace of instruction was steady yet it was not felt as pressure to finish up. Time was regarded as precious and was carefully budgeted, with the teacher pushing each child to do his or her best.

GENERALIZATIONS ABOUT THE MEDICINE CLASSROOM

This teacher approached most lessons mechanically, with little animation or personalization of the material. She openly read the guide in wooden, emotionless fashion, following the suggested strategy to the letter but in a perfunctory manner. She was not hostile to her students, merely distant. She seemed worried about keeping things going smoothly, keeping the lid on. Restraint, quiet, attention to the teacher and to directions were more important than involvement and interest in ideas.

The pace of instruction was slow, with many "empty" periods of waiting. The teacher's attention was focused on discipline in order to drag students through each "dose" of instruction. The "medicine" would do the job, it seemed, if the teacher insisted that everyone pay attention and go through the motions.

FURTHER COMPARISONS

Based on observations of both classrooms, observers agreed that "Sugar" was able to discipline the class, or individuals, without stopping the whole forward momentum of the lesson. The discipline "Medicine" dished out seemed to bring things to a halt. She couldn't seem to sort out extraneous noise from "learning" noise.

"Sugar" moved right along, rather quickly. She made a point to check every child's work. If "Medicine" did this, it was not obvious. The pace dragged in her classroom.

A big difference between the classes is that "Medicine" read the teacher's manual word-for-word to the class. "Sugar" just referred to the manual occasionally. She used most of the language in the manual, but it came out as her own, not as a recitation.

"Sugar" saw school time as time to stress learning; she expected students to do a certain amount of thinking. She communicated the idea that school is the only time students will get a chance to learn skills, so that time is valuable, and should not be wasted. "Medicine" saw school time as time to stress behavior and self-control. She communicated the judgment that her students won't get anything right unless they listen very carefully to her and do exactly what she tells them.

"Sugar" appeared warm and self-assured. She was seen as sincerely trying to prepare her students to be successful. "Medicine" appeared distant and aloof. She also appeared defensive or defeated at alternate moments.

Are these extreme cases? Somewhat. Many conditions have been ignored in this description that affect the teaching-learning situation. Probably as large a factor as the teacher's behavior is the student population and its characteristics. The focus here, however, is on instruction. How do these cases represent the field-test classes?

Looking at classes involved in the field tests, the proportion of classes with a medicine flavor is roughly twice as high as those that approach a sugar climate. Of course this is a gross oversimplification.

Some teachers were flippant and careless in their presentation. They were generally unprepared, turning to the activity for the first time at the beginning of the period and becoming frantic and disorganized when things were required that could not be immediately located or that required advance preparation. They tended to feel the materials should carry the whole load rather than serve as starting points, to be adapted to their class and pursued in the directions indicated by student response. Other teachers were never clear about where they were heading with their students. Their approach was academic, inconsistent, unresponsive, and often irrelevant. Still other teachers had no realistic idea of what was appropriate for various students and what were reasonable expectations to hold for them. Often these teachers knew very little about the student and seemed able to report only negative or damaging things about them. Other teachers were so involved in their own concerns (control, the subject matter, their teaching skills, etc.) that they were not yet ready to relate to and deal with students as people. Teachers representing the "Sugar" classroom were not at all alike, but as a group they were warm and responsive to students. They had definite standards and clearly communicated expectations for student performance. These were expressed through encouragement, praise, trust, and a valuing of the students' worth and usefulness. Absent was sarcasm and criticism. These teachers responded to the feelings of students and knew some of what was happening in their out-of-school lives. They created a sense of shared experiences in their relationships with students. ■

SCIENCE IN THE SUGAR CLASSROOM

CLIMATE

Entering this two-story, block-long building, one might suspect that it is about twenty years old, although it is well cared for. Doors are unlocked and students pass through the halls without being checked, even though they carry passes. No security guards are present. The test classroom was formerly a home economics room. It contains three sinks, a stove, a refrigerator, work tables, and movable desks and chairs for the students. The many counters and cabinets provide ample storage space. Resources and materials seem plentiful. Games and art materials are evident. The room is colorful and personalized with much student work. Some bulletin board areas appear to be designed by students. Others have activities or timetables designed for student use. The teacher's desk is at the side of the room. Student desks are organized in clusters oriented for group work. The following excerpts from several activities characterize the nature of classroom activities during science period.

Classroom Transactions: Presentation
UNIT I, ACTIVITY 1, "A Recycled Spaceship"

(From Observer's Notes) Students arrive from lunch. During loudspeaker interruptions, students engage in horseplay. The teacher passes out materials and moves students from desks to table groups. After almost ten minutes of housekeeping details she says, "If you want to know what we're going to do, you'll have to look here and listen." Students respond fairly rapidly and quiet down while directions are given following the teaching strategy in the manual. Statements from the manual are presented naturally, as if they are the teacher's own words. (Students are to build a model spaceship out of used materials: paper cups, scraps of paper, etc.) Two tables of girls get quietly but busily to work. Two tables of boys get to work with some horseplay, and there is much conversation about what they are doing.

Dave asks, "Is this art?" He stacks his cups one inside the other and blows Sandy's model over. He says he will bring more cups and make his model tomorrow, and then he begins a discussion of basketball with Sandy. The teacher is working most closely with another group. Mac is slow to begin. In fifteen minutes he has put two cups together. For most of the period he seems totally disinterested. He finally starts decorating his model and then plays with it, bouncing it in the air.

The teacher circulates, answering questions, making suggestions, and giving encouragement. After about five minutes, Dave begins to make a model spaceship. For thirty minutes students continue working. Some announce they are through in half this time, but they then start back to work adding to or decorating their models. Sandy is finished and is helping Dave. The teacher says, "We'll take about ten more minutes." Students groan and the teacher explains, "Ten minutes is a pretty long time." Students who have finished are directed to start cleaning up. Some horseplay occurs as models are completed, and the teacher attempts to begin a discussion, following (but not reading from) the manual. The students provide appropriate answers to questions. The group begins to get restless after ten minutes of discussion, but the teacher continues for five more minutes before ending the period. Students have shown enthusiasm and creativity in the activity.

(Teacher's response to question, "How did the activity go?" on feedback form) "They enjoyed constructing the models—a little slow starting but after encouragement did fine. Towards the end of the question period students seemed restless and knew answers almost too well. They enjoyed comparing real and model spaceships."

(Continued on page 47)

SCIENCE IN THE MEDICINE CLASSROOM

CLIMATE

Trying to enter this three-story, ornately styled building designed forty to fifty years ago, one finds all the doors locked but the one that provides access to the office. It is explained that this is done to "keep outsiders out." Security guards are placed in the halls throughout the building at strategic points. Students must have their hall pass available to be checked. Students are spoken to much more harshly than they ever would be at the Sugar school. A hostile, suspicious attitude toward them seems to be present. The test classroom is well lighted and warm. The noisy heater makes temperature control (and hearing) difficult. Windows are heavily screened. The room, somewhat overcrowded with furniture, has movable desks, few tables, and some storage cabinets. Resources and materials appear inadequate; some books are inappropriate, and there are not enough of some workbooks to go around.

The room is sterile and impersonal. No student work is displayed, nor are bulletin boards designed to be used by students. Fall pictures and decorations have been put up. The teacher's desk is at the front of the room with all student desks oriented toward it. Three or four students are seated in corners of the room, separated from the group.

Classroom Transactions: Presentation
UNIT I, ACTIVITY 1, "A Recycled Spaceship"

(From Observer's Notes) Students arrive in class and immediately notice cups stacked in a rather unobvious place. One of them says, "They're for building rockets." Bell rings, and they are quiet and generally well behaved as teacher takes roll. The teacher begins by pointing to a globe and saying, "Who knows that this is?" After the students unanimously answer, "A globe," the teacher says, "Okay, today we're going to make rocketships like we were taking off from this globe." The class is reorganized into groups of two, and the teacher hands out cups to each student. A student hands out scissors and tape. Then the teacher says, "Okay, please pay attention," and she reads the instructions from the manual. Students ask several questions, such as how they are to do it and whether they have to use all the cups. All begin busily building the models. After several minutes of silence, the teacher says, "Think of how different you can make yours—remember that spaceships of the future may be of different shapes." In the next twenty minutes the students work with little direct attention from or involvement with the teacher. She speaks only when asked a question: for example, "Do we get to take them home?" "Later. We are going to use them some more." Her expression is neutral most of the time, and she offers little encouragement or praise to anyone. Part of this time she works on something at her desk. The students remain quite involved in their work and talk about their models. As they finish, they play with them and continue to add details and decorations. About fifteen minutes before the end of the period the teacher has a student pass the wastebasket, and she asks students to turn in their scissors and tape when they are finished. During the next five minutes she looks at a few of the models and smiles warmly at their creators. Then she sharply calls for attention and asks everyone to take their seats. She reads the discussion questions from the manual, with little expression or modulation, and follows the teaching strategy for the remaining ten minutes of the period. Students provide lots of responses and are involved and interested in the discussion, which is cut short by the bell amid groans of disappointment.

The next day the teacher begins the period with the last four questions asked the day before. Students participate

(Continued on page 47)

SUGAR CLASSROOM (Cont.)

During the course of this period, visitors have come to the room twice, the PA has interrupted class once, a student is gone from class for ten minutes on an errand, and two students are dismissed early to go to another class. Many students move about the room getting materials, flying their spaceships, and showing each other their models. They talk freely with each other. The teacher occasionally tells some students to sit down and scolds others for horseplay, such as setting the trash on the stove or throwing water on another student. Three students are absent for this activity.

Classroom Transactions: Student-Teacher Interaction
UNIT I, ACTIVITY 4, "Sounds From My Environment"

This activity utilized a prerecorded tape of twenty sounds with accompanying slides picturing the source of the sound. Students received a worksheet with the following format:

	Is the sound in your environment?		What is the sound?	Did you guess right?	
1	yes	no		yes	no
2	yes	no		yes	no
3	...	...	.....	...	...
4	...	...	.....	...	...
...	...	...	.....	...	...

(From Observer's Notes) The teacher distributes the work sheets and gives directions for listening to sounds and marking the worksheets.

Mac asks, "What is environment?" The teacher has students tell what was said about the word yesterday in class. Mac asks, "Does environment mean around the school?" The teacher says, "Everywhere in your environment." The first sound (wind) is played as students listen quietly. The teacher stops the sound, reviews how to use the worksheet, and then replays the sound. Following the strategy but not referring to the manual, the teacher asks how many think the sound is in their environment (all but one do) and what their guesses are. (Most guess wind.) Bob guesses, "Tornado." Glen says, "It's not an ordinary breeze." John says, "Wind is the same as a tornado." The teacher explains that they can count all those guesses as correct.

Sound 2 (storm) is played. Jerry guesses, "Flying saucer," Jim says, "Train." The teacher has to quiet the class, and then she repeats the questioning strategy. Don guesses, "Train"; Jane says, "Plane"; Jim, guessing again, says "Rocket"; Ricky says "Traffic." Dave says, "It doesn't make sense. Is it a real sound?" The teacher reassures them by saying, "Don't be afraid of being wrong. The worksheets are not going to be corrected or anything. We're just practicing listening to sounds." John says, "I know what it is—a volcano." When the slide is shown, Ricky says it is a storm. The teacher asks why, and Ricky points out the rainbow, the fog in the mountains, and the clouds. None guessed rain or storm, but they agree that it is part of their environment.

The sound of a cow brings laughter. Some say it is not in their environment and some say it is. A few remain uncommitted. The teacher stresses differences in individual environments.

At the eighth sound, the teacher asks how the worksheets are going. Students all say fine. Jerry says, "This is fun, man." Students are involved, and there are cheers when a slide shows that their guesses are accurate. The teacher continues to accept many answers. During the period there is much movement of students from seat to seat. Students have gotten noisy, and the teacher waits until they are quiet before beginning next sound. Sandy kicks John in the side for blowing his paper off the desk during this interlude. John cries. The teacher suggests that

(Continued on page 48)

MEDICINE CLASSROOM (Cont.)

actively in the discussion for the first fifteen minutes. The teacher follows the manual verbatim. Each question is asked, even if it has been answered by previous responses. The teacher continues the discussion for another twenty minutes. Students become more and more restless, but generally they still participate. Throughout the period the teacher interrupts discussion occasionally to loudly discipline individual students, such as, "Put that paper under your desk. Do you know where *under* is?" The discussion is ended ten minutes before the period is over, and students are given free time.

(Teacher response to question, "How did the activity go?" on feedback form) "It went very well except that the class was above the questions and the discussion. They knew all the answers, so discussion dragged and seemed boring to them. They liked making the spaceships. My students have a hard time paying attention when just discussing; it is my feeling that a worksheet to fill out would have kept their minds from wandering and kept them from talking and playing with their neighbors."

Classroom Transactions: Student-Teacher Interaction
UNIT I, ACTIVITY 4, "Sounds From My Environment"

(See Sugar classroom for description of this activity.) (From Observer's Notes) The teacher passes out the sounds worksheets as students quiet down. George asks, "Will you give us time to write it down? How will we know what the sound is?" To the last question the teacher replies, "I'll show you a picture of the sound on a slide." She gives an example of hearing Indian drums beating to explain how to fill out the worksheet. Then she begins to read from the guide. Answers suggested after playing the first sound are: hard wind, raining, monster, storm, train, windstorm.

The following is a sample of dialogue during this activity:

Teacher: "Everybody be quiet."
(Silence, listening to tape.)
Teacher: "Put down yes or no; is that in your environment? Marvin be quiet."
Student: "Raining"
Teacher: "It was raining, you thought it was?"
Teacher: "Okay, here's a picture of it."
Teacher: "Okay, here's the fifth sound. Paul, start with number five, please."
(Silence, listening to tape.)
Teacher: "What did you write down, Lisa?"

In an ensuing fifteen-minute sequence, the teacher spoke twelve times, the students twice. Most teacher comments were to discipline or give directions. Little discussion was engendered.

Marvin writes "train" as an answer for four of the sounds. The tape is being played very loud and perhaps he is mistaking the loud background noise for a train. Stella stands out in the accuracy of her answers. When others say "cars" and "horns," she says "traffic." Both Stella and George are interested in checking the number they get right at the end of the lesson. They remain seated and are tallying the score. The teacher emphasizes that any answer is okay, but yet speaks in terms of the "right" and the "wrong" answer. The result is a guessing game of right or wrong rather than an inquiring process. Interest seems to build as the sounds are played. When the bell rings about halfway through the sounds, there are sighs of regret. (This activity was continued for four more days.)

The second day of listening to sounds on the tape continues with a high level of excitement over identifying sounds. The teacher has reduced the tape volume to decrease the background hum, which Marvin had mistaken for a train

(Continued on page 48)

SUGAR CLASSROOM (Cont.)

Sandy apologize, so he does. The teacher gives John a tissue, and he sits with his head on the table as three more sounds are played and discussed. Both boys continue to participate actively. The period ends.

The same listening activity continues for twenty minutes the next day. Students listen intently and remain enthusiastic. Jerry and John complain that there are *only* twenty sounds.

Following the teaching strategy, the teacher has students review their answers, listing on the board those responses that they agree on and those they don't agree on. Then the teacher divides the class into three teams to record sounds from certain environments around the school. Each team is sent out on its own. Some must wait in class for their turn with the tape recorder. Each team is gone about five or ten minutes. When all have returned, the teacher asks one student to keep a list of the sounds identified. Students are intent on listening to the sounds and giving answers. One group is delighted when they find out their sounds have not recorded, and they are allowed to go out again.

On the following day the tape for this group is played and discussed. Then the tapes for the other two groups are replayed. Students again are intent on listening. They ask to do the sound-collecting activity again. They want to make motions and act out the sounds, too.

JERRY—A CASE STUDY IN STUDENT-TEACHER DYNAMICS

The following excerpts from observer's notes illustrate Jerry's curiosity and interest in science and a teacher-student relationship that allowed for inquiry and exploration, in spite of some misunderstandings and interference.

UNIT I, ACTIVITY 18, "A Model Of My Neighborhood"

As students begin to make a map, the teacher wants them to include landmarks in the area. She asks, "What are some buildings or objects that are found in this area?" Students list the high school, the golf course, and some businesses. Jerry says, "How about Woolco?" but he doesn't say it loud enough for the teacher to hear. Later he is heard when he suggests the catwalk. Nobody seems to know what he is talking about. The teacher notices he is chewing on something and asks him to empty his mouth. He says no, and she says in a loud voice, "Right now!" He does.

The next day in the same activity, the teacher reviews the term "landmark" and asks students to list the most important ones they talked about yesterday. Jerry interrupts in a tone that indicates he is pretty put off. He says firmly, "All of us who ride the bus pass the catwalk.' The teacher asks him to describe what he means. When he does, it becomes apparent that he is referring to the overpass. The class generally agrees that this is an important landmark, so it is included in the list. Thus Jerry's two days of being misunderstood and rejected end on a note of triumph.

UNIT I, ACTIVITY 22, "Plants Need Water"

Students are told they are going to do an experiment with pea seeds. They place them in a test tube full of water and mark the water level. Jerry is fascinated by the way the peas wrinkle. He says, "There are two little eyes in there and a nose." The next day he is the first to notice that the water level has gone down. For his experiment, to see if pea seeds need water to grow, he placed some seeds on wet toweling. The teacher used Jerry's plants for the class to see where the roots come out of the peas (although all do an experiment). Sheryl says in an incredulous tone, "They come out of the peas?"

(Continued on page 49)

MEDICINE CLASSROOM (Cont.)

yesterday. She especially picks out Marvin to tell him that the background might sound like a train; "Understand, Marvin, *no* trains." After the first sound (sawing), the teacher again pointedly asks, "It didn't sound like a train, Marvin?"

There is general excitement now over each new sound and usually a cheer about the slide identifying the sound. Most animated are Betsy, George, Ann, and Carl. Leon is somewhat less excited. Least excited and involved, in fact almost dragging and depressed (at times), are Paul and Marvin. In both cases they began the period with a modicum of positive participation (laughs and groans at the answers), but they become more quiet as their wrong answers mount (case of Paul especially) or as the teacher disciplines (case of Marvin especially). Paul seems to persecute himself: after missing two in a row, he stuck to "motorboat" when everyone knew the next one was a "toilet." When he missed the next sound he threw his paper to the floor (worksheet).

The teacher continues to deny that any given answer matters, but continues to speak in terms of "right" and "wrong" answers, as in the following: "It had to be a plane door...A truck door is wrong." This statement was elicited partly in response to George's attempt to erase and change his answer. The question of changing answers and "cheating" added to the "right" and "wrong" viewpoint.

The third period on this activity is spent listing sounds on the board, with students very restless and inattentive. In six minutes thirteen sounds are listed on the board, by number, and the teacher has made three disciplining statements to students. This listing work is slow going. After the first ten minutes of the exercise, half of the thirteen students are showing attention to the teacher. The others are doing such things as writing, playing with a spaceship model, looking at the floor, etc. After twenty minutes of class, there are twenty-seven sounds listed. Only one-third of the students show any signs of being attentive. The teacher makes more disciplining statements: "Marvin, use only one chair!" (He was sitting backward on his own, with head on back of other chair, looking at floor.) "George, keep it quiet." (Room is getting noisy.) After writing sound thirty-three on the board, the teacher announces what they will be doing today—taping sounds outdoors, "But we won't get to do it if you keep talking..."

The teacher is stalling and apparently does not want to go out. The listing activity continues for another ten minutes, after which there is just time to review the running of the tape recorder before the period ends.

After class the teacher says to the observer, "From now on, when I do this, I'm going to make everybody *copy* in order to keep quiet." Somehow, the lack of interest is blamed on idle hands, ignoring the possibility that the material, or the way it was presented, might be just plain dull.

The fourth period begins with the class divided into three groups for recording sounds around the school. This period is full of fun and excitement. The teacher arranges for an adult to be with two of the three groups. The third group stays in the room to record. Before the groups fan out, the teacher reads word-for-word from the manual the instructions to the class. There is only one statement for the teacher to make in the day's teaching strategy.

There is a high level of excitement throughout the sound-collecting. Toward the end of the period Jose asks hopefully, "Are we going to continue this next period?" Students chorus, "Ohhh," when the bell rings.

The fifth and last day of this activity continues with a high level of excitement over identifying other groups' sounds. The period begins with the tape recorder on the teacher's desk and a stool beside the desk for the group

(Continued on page 49)

SUGAR CLASSROOM (Cont.)

UNIT I, ACTIVITY 23, "Plants Need Air"

Students have done another experiment, this time sealing some plants in a jar. After a week, the teacher has reviewed the experiment and has had them observe their results. She asks, "How do you explain the results?" Most plants without air (covered with metal lids) are not growing. Some plants with air have not grown. Sheryl admits she forgot to water them. Some plants covered with Saran Wrap have grown. Students notice this and decide that it must not be air tight. The whole discussion is conducted in such a way that conclusions are honest and not forced, they follow easily from what was observed. The teacher gives directions for disposing of experiments, but some, including Jerry, want to keep their plants. Jerry says earnestly in an undertone, "I wish I knew how seeds grow." Later he asks aloud, "Mrs. Sugar, how do seeds grow?" Ricky attempts an explanation, but it is obviously inadequate for the level of Jerry's curiosity. During a break, Jerry and the teacher converse about seeds.

A week later, Jerry proudly calls the observer's attention to his peas, which are growing well. ■

BEST COPY AVAILABLE

MEDICINE CLASSROOM (Cont.)

"leader" to sit on while his group's sounds are played. There is keen attention. George moves up two seats to hear better. Soon other students move to the front of the room around the recorder. The teacher at first asks, "Everybody please sit down." Soon they are all crowded around the recorder again. The teacher takes several deep breaths, as if to order everyone to sit down in their own seats again, but she appears to reconsider. She allows them to remain in front.

During the playback of sounds the teacher does not explain to students what they are doing or what they might learn from this. Some sounds that were recorded in the room are: stamping, pencil sharpening, clapping hands, snapping fingers, chewing gum, clapping erasers. Some of these produce much laughter. There is excitement over identifying any sound, even the easy ones.

JOSE—A CASE STUDY IN STUDENT-TEACHER DYNAMICS

As the following excerpts from observer's notes illustrate, Jose demonstrates an understanding, an insight, and an involvement in science that are repeatedly squelched by the teacher, who has rated him as naive and low in ability.

UNIT I, ACTIVITY 6, "Grouping Objects From My Environment"

There was considerable enthusiasm by the students for the activity, but it was dampened by the teacher's continued emphasis on discipline. Finally she shouted loudly, "JOSE!" who responded, "Can't I even talk?" (He was talking about the grouping process). The teacher wanted him and others to be listening at that time to another group.

UNIT I, ACTIVITY 10, "Temperatures In My Environment"

When it was time to hand in thermometers Jose was holding his, and he didn't pass it in right away. The teacher scolded him for this, saying, "Jose, now turn that thermometer in. I've asked three times..." He responded, "I'm waiting for it to get hot." (He was holding it in his hand, having just come in from outdoors, and watching to see if it would go back up.)

Later in the same activity...

Lisa, Stella, and a few others thought the temperature would be higher under the black cloth, because the black material was "thicker" and "heavier." George and the others thought it was hotter under the white, but they were not sure why. About four thought the temperature under both colors would be the same because cloths are the "same material" and "came from the same factory." Jose then asked, "Can we feel the cloth?" Teacher said "certainly," and Jose went up to do so. He then decided on black for sure, from the apparent heaviness of it.

UNIT I, ACTIVITY 11

The strategy: "I want you to measure how wide your desk is without using a ruler." Students should have invented a unit of measure and then measured their desks. (Most estimated in terms of number of rulers—"I say two rulers"—or number of inches.) The strategy suggests that if at least one of the students is successful in inventing a unit of measure, the teacher should use his idea.

Just as the teacher was about to show students how to measure with hands, Jose piped up and pointed out to the teacher that she had forgotten to ask him how he measured. She did, and he showed how to use a pencil. His desk was two pencils by three pencils. Instead of using his idea, the teacher went right on to illustrate measuring with hand widths.

After class the observer complimented Jose by saying, "You are doing well; you were the one to use a pencil to measure." The teacher then added, "But what he *should* do is remember to bring a pencil to Mrs. Johnson's class." ■

judgments–data payoffs and tradeoffs

Curriculum evaluation should inform developers of the strengths and weaknesses of the material produced, which in turn should lead to revisions that improve it.

Curriculum evaluation should validate the integrity of the material produced, providing evidence that it works in the classroom.

Curriculum evaluation should prove effectiveness, establishing how much the material significantly increases learning in the intended areas compared to other materials produced for the same purpose.

Curriculum evaluation should inform consumers, identifying the appropriate uses and limitations of the materials produced, thus leading to enlightened buying practices and proper utilization of the materials.

Curriculum evaluation should provide information to decision makers, supplying those who control policy and resources with data that leads to rational and informed planning of future efforts.

Each of these admonitions has its champions, who seem interested only in the data that they themselves need; efforts that provide other data are even criticized. Each represents a different audience: producer, teacher, school board, funding agency, legislator. An evaluator cannot serve two -or five masters well. He will exhaust his resources for the one and neglect the others, or he will be convincing to the one and questioned by the others. The results of an evaluation, however, often do speak to more than one audience. Such is the case in the evaluation of ME AND MY ENVIRONMENT.

At the highest level, that of the policy maker, the case of the two classrooms in the previous article raises a fundamental question: What is the payoff of continued expensive funding of curriculum development? The products provide the good teacher with vehicles that assist the educative process. They provide some refreshment, perhaps, to the students of poor teachers. But do they affect and improve the quality of the educative process? Mounting evidence from studies and reviews of the era of large-scale curriculum studies suggest that their impact on instructional practices has been less than revolutionary. When the evidence of the preceding article is combined with other studies, such as those of Gallagher,[1] what are the implications? It would seem that before further curriculum development efforts, an important avenue to improve the educative process lies in programs of teacher selection, training, and supervision. Both are needed, but curricula can have only limited success without training. It would also appear that the training needed is not merely programs to facilitate implementation of new curricula. The changes in teacher behavior are more basic than appropriate utilization of materials. They concern appropriate relationships with people (students).

The developer is usually not at liberty to act on information at this level. His mandate in most cases is very narrowly circumscribed by the funding agency.

The results speak to other audiences. To the teacher, there is some evidence to say that the activities work if used as directed. There is also evidence to suggest that serious distortions result from inappropriate emphasis. For the major aims of the curriculum to be realized, intellectual honesty to the processes of inquiry and problem-solving that are embedded in the materials seems necessary, as well as some fidelity to the intent of the individual activities.

The issue of effectiveness of the materials will be addressed to some extent in succeeding evaluation reports. Only limited results of this kind will be available. An evaluation of this type has not been funded. It should be

[1] *James J. Gallagher. 1966.* Teacher Variation in Concept Presentation in BSCS Curriculum Programs. *Urbana: University of Illinois, Institute for Research on Exceptional Children.*

conducted by someone other than the developer, using the final product rather than experimental materials. The same study could provide consumer information.

PAYOFFS

Turning to the central purpose of the present evaluation—providing information to guide revision—what were the tradeoffs and payoffs of the forms of data used?

Counting the costs and placing a value on forms of data are difficult tasks. The costs do not all have a monetary value. Some kinds of data represent a heavy burden on the contributors, or the processors. Some take a toll in attitudes toward the study. Some are costly in the distortions they create. The usefulness of the data is one value that will be estimated. This judgment gets mixed up with the nature of the results themselves, whether they were positive or negative, and their importance or significance as findings in themselves. Forms of data that were "merely" confirming, without adding much new information, tend to be undervalued for the credibility they establish in the results. It is also almost impossible to determine how much certain kinds and pieces of data have influenced the staff, although data sources that seemed to make a big impression on the staff will be identified.

Considering first the six major sources of information, by far the most useful data for revision came from the teacher. The utility of this information, however, is derived from having first carefully selected participating teachers and then having identified which of those were objective, insightful, and creative in the generation of alternatives and revision suggestions.

Second in its impact and usefulness for revision were detailed reviews of the content by specialists and staff. Part of this payoff lay in the caliber and diversity of the reviewers, part lay in their specific comments and the effort they invested in suggesting additions and alternatives. The payoff in staff review came in the repeated matrix analyses and in the care with which the materials were edited and made consistent in format, style, and sequential development.

Third in utility for revision was data obtained by the staff's direct observation of instruction. This source of information had by far the highest impact on subsequent directions to revision writers. An important element in the influence of this source was in producing written reports of each observation. The hidden component was how much these limited views influenced each staff member's interpretations and recommendations of all data. It also served to identify teachers whose perception and understanding was judged exceptional. The combination of teacher and staff feedback was deliberately structured to have the greatest amount of impact on writers involved in revision.

Fourth in order of immediate usefulness for revision was data on student abilities and performance. Over the course of several years this information has influenced the materials at a general level and has led to the introduction of a number of activities that would not otherwise have been included.

The use of observers other than full-time staff did not have a consistent utility for revision. As a format was developed for what to observe and how to report it, however, the usefulness of this information to writers was increased. Greater value and import could have been achieved with a greater monetary input into the further development of observation protocol systems or alternatives. Use of ethnographic techniques would also have increased the usefulness of this source of data. In spite of this low rating in relation to other sources, the value of observers' reports must be judged high for two services they performed. One payoff was the verification of teacher feedback, or the identification of poor questions that were impossible for teachers to answer meaningfully. A second payoff was the occasional identification of activities where there was a problem, when teacher feedback contained no hint of it. Only through an observer's view were a number of mistakes in instructions and strategies recognized. While such findings were somewhat infrequent, failure to rectify them would have weakened the materials.

School records were of no utility for revision purposes. It should be noted that this ranking of utility and impact refers strictly to revision uses of the data. There were other quite valuable payoffs. For example, data on student abilities and performance, ranked fourth above, has led to an increased understanding of the retarded child that will definitely influence future development efforts. Even the collection of school records had a payoff in the negative sense: it established that these records have no value for guiding revision or explaining variance in performance.

SPECIFIC JUDGMENTS

Referring to Table 2 on page 32 listing data collected from field-test classes, the following judgments can be made.

1. Demographic data on students (Items 1-4 on this table) have had absolutely no value for this evaluation. They suggested nothing to writers, they explained nothing in performance, they had no instructional utility.
2. General descriptions of students (Item 6) suggested some functional characteristics to explore in the early development of materials. They also provided a perspective of the student population that broadened the view obtained from site visits.
3. Teacher ratings of students (Items 7-14) were obtained in an exploration of characteristics and functional abilities that might relate to performance. Some dimensions (Items 10, 13) were identified that did

(Continued)

have explanatory power. (See Reports 2 and 4.) Additional ratings (Items 37-43) were introduced as components of the program.

4. Measures of cognitive development and problem-solving ability (Items 15 and 16) have been developed that proved to be far more meaningful than IC. They also have the potential for setting realistic expectations of students and guiding instructional decisions. These measures have had a general influence on the overall design of materials.
5. Measures of entering ability of students (Item 17) proved to have high utility for revision. They have also served a useful function for teachers, and have been included as a component of the materials.
6. Measures of performance (Item 18) were used to judge the efficacy of sequences of instruction. They have influenced revision considerably because of a requirement the staff made that they have utility to the teacher and be included as a component of the materials. The most meaningful of these assessments have involved situational tasks requiring an application of a skill or idea.
7. Obtaining completed student worksheets (Item 19) has been of direct utility for revision. These have been most helpful when coupled with teacher or student explanations of problems.
8. No accurate assessment of individual student interest (Item 20) has been developed, but ratings of group interest and involvement have influenced the judgment of each activity's effectiveness. Low ratings in this dimension have invariably led to revision.
9. A record of absences has had no utility in revision efforts.
10. Ratings of lack of success did not reveal any activities that needed revision except those with already highly visible defects. In each classroom there tended to be a student or two who functioned far lower than the rest of the class. Teachers usually named these students time after time.
11. Validity studies, such as Item 24, have had a high utility and generally provided information useful for revision in addition to qualifying other data.
12. Attitude questionnaires (Items 25-29) have proved helpful in identifying which activities in a sequence need most attention. Their use by revision writers was similar to the use of the interest rating.
13. Samples of various abilities that were subjects of instruction (Items 30-32) served to adjust the degree of emphasis and development of these abilities in the materials. They were, however, peripheral to the main revision effort.

TRADEOFFS

Traded off in the design of this evaluation and choice of data to collect were two things: 1) breadth of understanding that a closer study of the complexities of the classroom might have provided; 2) exploration of individual learning problems that longitudinal studies of individual children would have provided. Both of these efforts would require major investments. Their payoffs would be slower, longer range, but perhaps ultimately more basic. The choice of data to collect and the design for collecting it depend upon the level of generality and the breadth of information desired. Developmental efforts must trade off generality for an abundance of specific details. For a group-centered program such as ME AND MY ENVIRONMENT, trial in one classroom would provide an indication of how students respond to the materials. It would also be possible to note a number of details directly and follow individual students closely, for attention would be narrowly focused on one time and place. The use of video and audio tapes and much informal interviewing would increase the richness of the data. This represents a considerable added investment generally not available unless the sample studied is strictly limited. Another perspective could be gained by arranging for staff to teach a class of children. Reduction and transmission of results could be further simplified by arranging for revision writers to participate in such field trials. Given the experience gained in the current evaluation, this is the direction the staff would recommend for future studies.

At the time this project began, one concern was to be sure the observed responses were typical and that the curriculum would function under a variety of conditions. Intensity of study was traded off for this assurance. An attempt was made to maintain some degree of direct involvement and contact with test sites, however many judgments and interpretations were made at a distance. To compensate, redundancy was built into the data collection scheme. The plan has worked. A considerable amount of data has been available and utilized in revision decisions.

In the end one must recognize that there is no best way to evaluate, no ready answer to such questions as "How many test sites are enough?" and "What data shall I collect?" Instead the choice must represent a comfortable fit with the purpose, and there must be a recognition of what has been lost as well as gained by the choice.

The case study in this section was an attempt to study this tradeoff of portrayal for representativeness. It is one we will not as willingly make the next time. But the alternative decision, too, entails a tradeoff. There are never enough resources and time to answer all our questions. In evaluation we are left, like Jerry in the Sugar classroom saying wistfully, "I wish I knew how these [curriculum] seeds grow." ■

BIOLOGICAL SCIENCES CURRICULUM STUDY

Me and My Environment

ACKNOWLEDGMENTS

The development of a curriculum is the product of many people. We are indebted to the five members of our advisory board for wise counsel in the design and planning of the program. From seven reviewers came substantive ideas for modification. A number of evaluators contributed to the plans for data collection and analysis. The field test teachers did yeoman service in providing the bulk of data for study. The insights of six observers further contributed to the richness of data collected. The analysis of materials and data from field tests has represented a joint effort of the staff consultants for the project. The contributions of the illustrations staff to every phase of program development goes often unrecognized, but represents a significant increase in the meaningfulness of all materials.

Thanks are also due to Melissa Cooney for the care with which she composed final camera copy for printing.

ADVISORY COMMITTEE

WILLIAM R. REID, *Chairman, University of Florida, Gainesville, Florida;* **ROBERT L. ERDMAN**, *University of Utah, Salt Lake City, Utah,* **WARREN E. HEISS**, *Montclair State College, Upper Montclair, New Jersey;* **JERRY H. HUBSCHMAN**, *Wright State University, Dayton, Ohio;* **EDWARD L. MEYEN**, *University of Kansas, Lawrence, Kansas.*

EVALUATION CONSULTANTS

PHILIP BASHOOK, *University of Illinois, Medical Center, Chicago, Illinois;* **AUSTIN J. CONOLLY**, *University of Missouri, Columbia, Missouri.* **ROBERT ENNIS**, *University of Illinois, Urbana, Illinois;* **GENE V. GLASS**, *University of Colorado, Boulder, Colorado;* **WILLIAM M. GRAY**, *University of Dayton, Dayton, Ohio;* **ERNEST R. HOUSE**, *University of Illinois, Urbana, Illinois;* **JAMES A. JONES**, *Syracuse University, Syracuse, New York;* **WILLARD JONES**, *Rocky Mountain Special Education Instructional Materials Center, Greeley, Colorado;* **OLIVER KOLSTOE**, *University of Northern Colorado, Greeley, Colorado;* **EDWARD KORMONDY**, *The Evergreen State College, Olympia, Washington;* **MARLYS MITCHELL**, *University of North Carolina, Chapel Hill, North Carolina;* **BARAK ROSENSHINE**, *University of Illinois, Urbana, Illinois;* **DOUGLAS SJOGREN**, *Colorado State University, Fort Collins, Colorado;* **JAMES WHORTON**, *University of Florida, Gainesville, Florida.*

1971 SUMMER WRITING TEAM

THOMAS E. ALLEN, *Florida State University, Tallahassee, Florida;* **JOSEPHINE A. BENNETT**, *Whitehaven High School, South Memphis, Tennessee;* **DONALD R. DAUGS**, *University of Victoria, Victoria, B.C., Canada;* **FLORENCE B. HAAG**, *Ida Patterson School, Eugene, Oregon;* **CREGG INGRAM**, *Southwest Educational Development Center, Cedar City, Utah;* **JO ANN JEPPSON**, *Granite Park Jr. High School, Salt Lake City, Utah;* **KEITH MORRILL**, *South Dakota State University, Brookings, South Dakota;* **NORRIS M. ROSS, JR.**, *Nicolet High School, Milwaukee, Wisconsin;* **WILLIAM G. VOLANTE**, *Turkey Hill Middle School, Lunenburg, Massachusetts;* **MARTHA M. WILLIAMS**, *Kepner Jr. High School, Denver, Colorado.*

1972 SUMMER WRITING TEAM

WALTER AUFFENBERG, *Florida State Museum, Gainesville, Florida;* **WILLIAM EDWARDS**, *Laramie County Community College, Cheyenne, Wyoming;* **WREN GREEN**, *University of British Columbia, Vancouver, B. C., Canada;* **FLORENCE HAAG**, *Ida Patterson School, Eugene, Oregon;* **JERRY HUBSCHMAN**, *Wright State University, Dayton, Ohio;* **CANDACE LIGHT**, *Hodgkins Jr. High School, Westminster, Colorado;* **KEITH MORRILL**, *South Dakota State University, Brookings, South Dakota;* **ALICE OGURA**, *Balarat Environmental Center, Denver, Colorado;* **JOAN PETERS**, *Boulder, Colorado;* **NORRIS M. ROSS, JR.**, *Nicolet High School, Milwaukee, Wisconsin.*

1973 SUMMER WRITING TEAM

JUNE E. ALLBRANDT, *Lake Jr. High School, Denver, Colorado;* **LOUIS R. CLERICO**, *Hillside School, Montclair, New Jersey;* **STELLA GUBBINS**, *Boulder, Colorado;* **FLORENCE B. HAAG**, *Ida Patterson School, Eugene, Oregon;* **WARREN E. HEISS**, *Montclair State College, Upper Montclair, New Jersey;* **JERRY H. HUBSCHMAN**, *Wright State University, Dayton, Ohio;* **CANDACE J. LIGHT**, *Hodgkins Jr. High School, Westminster, Colorado;* **KEITH MORRILL**, *South Dakota State University, Brookings, South Dakota,* **GEORGE E. NOWAK**, *Madison High School, San Diego, California;* **MARY D. SMITH**, *Shepard Jr. High School, Durham, North Carolina.*